Hans Jürgen Holstein

HOMO CYBERNETICUS 4

Artificial Psychology and Generative Micro-Sociology

UPPSALA: SOCIOGRAFICA. 1974

Acta Sociocybernetica Upsaliensis

Digitized by the Internet Archive
in 2026

https://archive.org/details/IA41548924_0010

Hans Jürgen Holstein

HOMO CYBERNETICUS

VOLUME FOUR

Artificial Psychology and Generative Micro-Sociology

UPPSALA: Sociographica, 1974

UNIV.
LEIDEN
BIBL.

Copyright © by Hans Jürgen Holstein

Preface

Homo Cyberneticus comprises five volumes. All are relatively self-contained, but complement each other into one whole. The *first* volume is a synopsis of the simulation intentions and approach adopted by us; its tacit main question is: "How can we operationalize these intangibilities?" The *second* volume tackles this problem on the methodological level by trying to elaborate a systematic epistemology for dynamic causal modelling. Its conclusions support our initial 'hunch' that simulation is the optimal tool for dynamic causal modelling. Volume *Three* is therefore an investigation into our possibilities of constructing a simulation tool (a computer program) which will allow us to concentrate on our problem instead of on the computer. We do not wish to adapt our problem to the computer, but expect the computer to adapt to our problem instead, so to speak. Volume *Four* leaves all computer problems behind and asks again how we can simulate human behavior; the investigation is this time less superficial than in the explorative first volume. The *fifth* volume, finally, delineates a framework for rendering the causal simulation of societal macro-processes possible. It is simultaneously a first example of how the principles discussed in the earlier volumes can constrain, generate, and guide large-model concretizations.

The tackled task is vast, and although everyone agrees that it needs be tackled, there seems to be a general suspicion that those who do it are charlatans and/or self-appointed prophets. It has hitherto been difficult to argue for *Homo Cyberneticus* properly because its 'point' lies in the linking together of a plurality of diverse fields of research,

and also because it contains relatively much 'New Think'. Especially the advocation of common-sense thinking at the expense of statistics has evoked skepsis and territoriality defenses. Sociologists have thus downgraded it all to 'mere' computer science, while computer scientists have considered the simulation project oddly humanistic and too problem-oriented.

I was therefore very grateful when the Swedish Riksbank Tercentenary Fund dared to given financial support to my research in 1968. It has done so until this day.

The other important event was the foundation of the Computer Science Laboratory in Uppsala by Dr. Erik Sandewall and myself, as a 'home' for our up to then 'homeless' research projects, his being in the field of artificial intelligence, mine in socio-cybernetics. This symbiosis has, at least for my part, been immensely stimulating.

I Thank also my collegues *Lennart Stålberg*, *Torgny Tholerus*, *Tore Risch*, and *René Reboh*, most sincerely for their participation and cooperation in this sociocybernetic venture during the past years.

Uppsala, in April 1974 H.J.H.

Table of Contents

PART I: *Artificial Psychology*, 1

1. Introduction, 3

1.1. Short Overview of the Robot-Psychology Field, 5

2. 'Abstract Biology', 8

2.1. Why do Organisms Behave at all?, 9

2.2. Basic Biological Mechanisms, 10
2.2.1. Generalities Regarding Anatomy and Physiology, 12

2.3. Needs and Behavior, 15

2.4. Personality and Behavior: Two Description Levels, 18

2.5. Accomodation, 21
2.5.1. Attention and Vigilance, 21
2.5.2. Conspicuity and Memorization, 22

3. Learning, 24

3.1. The Learning-Theory Labyrinth, 26

3.2. Three Simple Models of Elementary Learning, 30

4. Pattern Recognition, 35

4.1. Trainable Pattern-Classifiers, 36

4.2. Apperception, 38

4.3. A Model of an Apperceptor, 39
4.3.1. Example: Unambiguous Training, 40
4.3.2. Example: Ambiguous Training, 43
4.3.3. Example: Inverted Use, 45

4.4. Comment, 46
4.4.1. Inductors and Deductors, 46
4.4.2. Concepts and Conceptions, 47

5. The Personality Network, 48

5.1. Thinking and Reasoning, 50

5.2. Problem Solving and Planning, 51

5.3. Conscious and Subconscious Thinking, 53
5.3.1. Concept Formation, 54

5.4. A Simple S-O-R Model, 56

6. Semantic Databases, 60

6.1. Semantic Descriptions, 62
6.1.1. Contextual Meaning, 63

6.2. Semantic Inferences, 64

6.3. Pragmatic Inferences, 65
6.3.1. Implicational Molecules, 66
6.3.2. Abilities and Requirements, 66

7. Emotion, 68

8. Volition, 70

8.1. Volition and Ambition, 73

8.2. Free Will, 74

9. Self-Awareness and Self Image, 75

PART II: *Generative Micro-Sociology*, 79

10. Schedules, 81

10.1. Life Phases and Development Schemes, 83

10.2. Short Example of a Development Scheme, 88

10.3. Daily Curricula, 93

11. Maps, Places, and Travels, 94

11.1. An Example: Travelling by Bus, 95

11.2. Geographical and Economic Information, 97

12. Social Contexts, 99

12.1. Social-Context Memberships, 100

12.2. Social Norms and Social Acceptance, 102

12.3. Social Maturation, 103

12.4. Acquaintances, 104

13. Imaginary Contexts, 105

13.1. Roles, 106
13.1.1. Role Genesis and Modification, 107

13.2. Mass Media, 108

14. Coalition Theory, 109

14.1. Solidarity, 111

14.2. Attractiveness, 113
14.2.1. Social Apperception, 113
14.2.2. Making Impression, 116

14.3. Social Power, 117
14.3.1. Compatibility, 120

15. Emergent Properties, 123

15.1. Attitudes, 124
15.1.1. Cognitive Dissonance, 124
15.1.2. Resistance to Attitude Change, 125
15.1.3. Freedom of Choice and Dissonance, 125

Appendix, 130

16. References, 131

PART I: ARTIFICIAL PSYCHOLOGY

1. Introduction

We shall in this text sometimes speak of 'robots'. The word "robot" can be regarded as being almost synonymous with "model of man" or "simulated person", but we wish to allude to the idea of comprehensiveness of the aspired model of man by using the word "robot". Simon's (1957) 'models of man' would thus not be 'robots' in our sense of the word, because they lack comprehensiveness (although not generality). Besides, very much of for us most relevant research has been made in connection with attempts to render real robots intelligent.

Please note that we have given higher priorities to comprehensiveness and coherence than to complete adequacy in model details. We have reasoned that the improvement of details would be easier to organize within the framework of a 'complete' and at least superficially adequate man model, and that choices between alternative models could be easier within a context.

The cognitive capacities of robots of today are still very limited. Robots are therefore, as yet, quite imperfect models of human beings in that respect. There are numerous simulation models of various specialized cognitive abilities such as recognition of handwritten numbers, chess playing, theorem proving, face recognition, analogy reasoning, action planning in a simple environment, concept formation, rule learning, natural-language understanding, and conversation, but there is no single cognition model that can do all of these things (cf. Uhr 1973, who seems to try to approach such an integration).

We believe that the construction of an integrated cognition model requires a new approach that is yet to be found. The perfect cognition model must be trainable from the start to learn new cognitive abilities, precisely like a child. That the training may take years to accomplish, does not matter; after all, humans need also quite long times to attain their cognitive abilities.

The social robots that we envisage in this booklet shall one day be able to go beyond advices and derive own behavior plans from their basic needs and dispositions and in relation to the environmental givens. They shall be able to *invent* new behavior patterns if none of the given ones seems adequate to them, and they shall thus be truly able to solve problems they encounter. They shall also be able to communicate with us in a natural language and will *understand* what we say and what they say. That is, they shall themselves '*reason*' semantically and pragmatically, not merely perform syntactical manipulations on commands and descriptions. How distant this future is depends perhaps somewhat upon how many people will be engaged in general artificial-intelligence research during the coming years around the world.

1.1. Short Overview of the Robot-Psychology Field

The crucial question is, in this area of research, all the time "How can we get these electronical devices behave as if they *understood* their own behavior and the behavior of their environment?" Let us assume that the robots we speak about are to cope with an environment similar to the everyday world we live in ourselves. That means that our robots must be equipped with a *perceptor* that can detect signals in the flood of existing structures and ongoing processes and interpret them. Within the perceptor we must distinguish the *sensory receptors* (input organs to 'nerve channels'), the *accomodators* (which filter received sensations such that only 'extraordinary' stimuli ('signals') are transmitted along the channels), and the *apperceptors* (which somehow interpret transmitted signals).

Receptors and accomodators seem to be relatively easy to model, but we know less about how apperceptors work. We assume that apperceptors are pattern-classifying devices which 'produce' the basic elements of reasoning. These basic elements (called 'appercepts' or 'basic concepts') are idealized representations of signals; similar signals have the same representation in the robot's database which we usually call 'memory'.

Concepts are either basic concepts or hierarchical compositions of basic concepts. *"Concept formation"* refers to the process of usually inductive composition of a hierarchy of concepts.

"Learning" is an ambiguous reference to various processes of database modification. (We ignore the usage "sensory learning" which we call 'accomodation')

"Learning" means most often what we shall call 'implication learning' or 'belief formation', namely the establishement of a 'link' between one concept and another. (We do not know how these 'links' are represented in the brain, and we do not need to know as long as we do not attempt simulation on the neurophysiological level. Simulation of learning and reasoning can, on the psychological description level, be quite satisfactorily performed in terms of databases and linguistic structures.) 'Implication learning' can be achieved *inductively* (that is, through observation of the relative frequency of cases where the implication holds), or *deductively* (that is, through the application of rules of logic). "Rule learning" is, in our opinion, synonymous with (inductive) "belief formation". The rules of logic are learned beliefs (cf. Volume 2, Part 3).

Language use involves the apperception of syntactic patterns (which is a translation into semantic structures) as well as the generation of syntactic patterns from semantic (conceptual) structures. The set of (learned) rules for how the first translation is to be performed, is called 'semantix' (and it includes the rules of logic); the set of (learned) rules for the second translation (language generation), is called 'syntax' or 'grammar'. 'Pragmatix' is a set of rules which assign factual-validity values to conceptual structures. (We use the words "pragmatics", "semantics" and "syntactics" to denote the *study* of such rule sets.)

'Planning' and (symbolic) 'problem solving' are (usually deductive) processes of concept combination according to semantix and pragmatix.

The 'embodiment' of a robot is a set of mechanisms. Some of these are rigid (the pain mechanism, for instance), others are trainable (the apperceptors for instance).

The robot's motives (needs, goals, plans) have priorities. These priorities are determined by the embodiment mechanisms. Some of these determinations will be subconscious, other ones conscious. Conscious determinations can override subconscious ones; a robot can thus subject himself to 'painful' influence because it is necessary in the persecution of a higher-priority plan; were there no such plan he would seek to avoid the 'painful' influence.

Thinking involves the composition of concepts (or beliefs). Thinking can be *subconscious* or *conscious*. Conscious thinking is called 'reasoning'. 'Subconscious thinking' is a fuzzy concept which refers to reactions in the embodiment's mechanisms on various levels of fixedness: the concept pair Conscious-subconscious alludes to a lattice (or even a continuum) of trainability and flexibility of embodiment mechanisms. (A tentative classification of everyday-language labels of embodimental responses according to this conception yields the sequence reflex (or instinct) - drive - affect (or feeling) - emotion - - concept - belief - attitude - wish - plan - - intention.) This lattice or hierarchy reflects a systemic composition of the embodiment; the elementary mechanisms are simple and rigid but can be coupled in various manners to other ones, to form higher-order (and more flexible) mechanisms.

2. 'Abstract Biology'

We shall in this section try to outline a theory of behavior which is less abstract than that of general automata theory, but still not as specific as biology or psychology. We shall call this theory 'abstract biology' because we want it to be general enough to be applicable not only to natural organisms but also to artificial ones.

2.1. Why Do Organisms Behave At All?

The phenomenon 'behavior' is usually explained either in terms of motivations or, more specifically in terms of (physiological) causes and effects: once the system of (direct and indirect) interactions of subsystems which compose the organism is started it will keep on to self-maintain itself as long as its impulse-carrying interconnections are intact.

Or, in general: An isolated set of interconnected systems s_i is at a state of equilibrium if and only if all the composing systems s_i are simultaneously at equilibrial states. (Ashby 1956, p 82.)

Since the organism usually during all its life is in contact with a fluctuating environment, it is hence constantly *forced* to react. First when it is dead can the composing subsystems go independently to equilibrial states: They cease to be interconnected. It is also possible to deduce that the *destruction* implies the arrival at an internal state where at least two of the composing subsystems are completely isolated from each other, and that the *death* of an adaptive system is its partitioning into non-adaptive components.

2.2. Basic Biological Mechanisms

The 'basic' biological mechanisms which we shall discuss now are the intraorganismic processes which bring about an organism's capability to behave at all and which 'trigger' adaptive behavior.

We conceive of an organism as a time-spatially and materially connected natural automaton. We now introduce the concept *external to the organism* as a primitive notion: we assume henceforth that we can determine with (the for whatever purpose we may have at the time of relevance) sufficient precision whether something is external to the organism or not. The organism itself is, of course, the only set of phenomena which are not external to it; these are called *internal*. Whenever we cannot determine reliably, we may say that the phenomenon is *neither* external *nor* internal or *both* external *and* internal.

The *environment* of an organism is the totality of phenomena which are external to it. The *immediate environment* of an organism is the totality of phenomena external to it which 'directly' influence the organism; the exact meaning of 'directly' is dependent upon the level of detail used in the description where this concept is used.

The organism and its environment constitute together a cybernetic system (of which they thus are subsystems). The organism and its environment may have relatively few elements in common. These are called *interface elements*. The set of interface elements is called the *boundary* between the two subsystems.

Organisms are complex *input-dependent* systems. Input-dependent systems cannot take arbitrary inputs; some

inputs will namely lead to disorganization by inducing local stabilities which *disconnect* the system into subsystems which are functionally not equivalent to the composite system. Such disconnections we shall call *injuries to the system.*

Under the assumption that injuries are to be avoided, are all input-dependent systems regulation - dependent. That implies: in an evolutionary context can only *adaptive* input-dependent systems survive. Adaptive input-dependent systems shall here be called *intrinsic regulators.*

A *hierarchic system* is a set of subsystems (on some nth level of resolution) which are connected in a hierarchic manner.

Injuries to top-hierarchy subsystems of a hierarchic input-dependent system are more *severe* than injuries to subsystems of lower position in the hierarchy. All injuries which render an intrinsic regulator less effective are severe in proportion to the decrease in effectivenesss they induce. More severe injuries often lead to more extensive disconnections than less severe ones; but the estimation of the severity of an injury must also be based upon the increase of proneness to injuries which is the consequence of an injury.

Injuries to an organism which have such effects that it disconnects into systems which all are of lower complexity are called *fatal*. The organism is said to be *dead* as soon as injuries to it prove to be *de facto* fatal.

Organisms are *multistable* intrinsic regulators. This means (a) that organisms comprise, as parts, regul-

ators which tend to reduce input-dependence, and (b) that organisms exert influence upon their environments in order to control influences upon themselves (that is: their embodiments and internal processes).

We assume organisms to be hierarchic systems.

2.2.1. Generalities Regarding Anatomy and Physiology

An organism's embodiment comprises (a) the organism's anatomy, and (b) the organism's physiology.

The *anatomy* of an organism is the material system which manifests the physiological processes and the behavior of the organism. (States of the organism must manifest themselves as material states; physiological processes must involve material 'carriers' of energy or of information.)

The *physiology* of an organism is the totality of its physical and chemical control subsystems (which are necessary in order to prevent the organism from death (disassemblage)).

Organisms are (or happen to be) constructed in such a way that injuries have as side-effects signals of control priority in proportion to the severity of the injury.

Explication: In cases where several control signals can arrive as simultaneous inputs to a single regulator which cannot 'execute' all of them simultaneously, the regulator ought to be able to order the signals according to the same criterium; intensity

is usually such a criterium. A signal s_1 is said to be of *higher control priority* than another signal, s_2, if the regulator orders s_1 such that it tends to be "executed" earlier than s_2.

Control signals induced by injuries (or at least: arising as side-effects of injuries) are called *pain signals* (on the biological level).

Assumption: Organisms are intrinsic regulators which are able to discern control priorities among pain signals.

The *behavior* of an organism is the synthesis of behaviors of its subsystems; since the concept 'behavior' arises as an abstraction from the biological level, it is a psychological concept.

Bodily subsystems of an organism may generate control signals when they attain certain states. These states are associated with increased probabilities for injuries.

An explanation why this happens to be the case can only be given in an evolutionary context: Systems without such warning signals have higher probabilities for failure; they will therefore either fail to reproduce (if they are natural) or not be reproduced (if they are technical). Mutations and redesigns may reduce the proneness to failure in new systems.

The distribution of signalling and non-signalling subsystems in organisms has an analogous rationale, as well as the distribution of associated control priorities.

These control signals we call *discomfort signals*.

Pain signals are discomfort signals of high control priority.

It is probably a pseudoproblem whether control signals *cause* or whether they only modulate (influence, codetermine) behavior, or both.

We assume therefore that subsystem behavior *implies* the continuous modulation of total-system behavior.

An organism's behavior may be said to be modulation of the environment.

2.3. Needs and Behavior

(a) We make the metaphysical hypothesis that all actions are reactions. This hypothesis amounts to the suggestion that the empirical reality is composed of an 'infinity' of small automata that change their manifest states only in case their inputs change. *Time* does not exist for these automata, only *past*, *present* and *next* as regards their states. (That means that for the organism there is no chronological time; time, in the human-less context, must be measured in terms of the occurrences of changes in subsystems of the system under inspection. The next-state function in an organism must be assumed to be 'in reality' a summary of a multitude of minute intra-organismic changes.Each change, in turn, is a reaction to a change in the respective environment. What we call *behavior* is thus a pattern of sequences of reactions.

(b) We obviously propose an 'ecological ' view upon an organism. It is made up of a large multitude of small, simple subsystems, each with only one or a few output states, and reacting upon only very few environmental conditions or changes.

We can imagine that increases and decreases in activity intensity manifested by the total organism is a reflection of the proportion of subsystems being in environmental conditions to which they are determined to react with an output change, which in turn may affect other subsystems, etc., in a chain-reactive manner.

When we wish to allude to this 'ecological' conception of the organism's embodiment we refer to it also as the organism's *internal environment* (in

contrast to the *external environment*).

(c) This suggests that *needs* in an organism can be explained on a level of greater detail in terms of selective inputs (which is a universal phenomenon, present in all empirical systems; no empirical system is influenced by all other empirical systems which exist). *Needs* are, from this 'ecological' point of view, *trends* of behavior in the population of subsystems of the investigated system.

(d) If there were no *needs* in an organism, it could have no *goals* or manifest intent-directed behavior or possibilities to distinguish between agreeable circumstances ('rewards') and disagreeable circumstances ('punishments'). This very distinction is of such utter importance for all adaptive (=regulating) systems that we can even claim that systems which cannot make such a distinction shall not be called adaptive.

It is here quite irrelevant that *we*, as observers, distinguish further between *intrinsic* needs or goals (for instance, physiological ones), and *extrinsic* (imposed) norms or goals (for instance, the desired temperature as a norm imposed by us upon a thermostat; or a social norm imposed upon people by society). Also imposed norms must be somehow *represented as intrinsic needs* in the adaptive system, else they cannot become effective as behavior-determining criteria.

To put it provocatively simple: The notions of *x is good* and *y is bad* must somehow be embedded in the embodiment of an adaptive system; it seems after our discussion, justified to speculate that the

abstract, general notions of 'good' and 'bad' have a multitude of concrete embodimental (for instance, physiological) correspondents.

(e) In biological and psychological texts one often encounters a distinction between *primary* and *secondary* needs or goals. The explicit distinction, however, is often very vague. We feel that this distinction is justified and shall specify it as follows: We say that a system's need or goal is *primary* or *basic* if it is a requirement which *must* be fulfilled if the system embodiment is not to be injured. We say that a system's goal or need is *secondary* or *derived* if it concerns the bringing about of a state with lower frequencies or probabilities of intensities of basic needs. (For instance, *eating* is a basic need and goal of humans; the accumulation of money and/or of power so that one can secure one's access to food is a secondary need and goal in human society members).

(f) *Adaptation* is a result of a system's adoption of new secondary goals in a given or changed environment. *Basic* needs or goals can never be changed by the system (as long as the embodiment remains unchanged); *secondary* goals must necessarily change if the environment changes in relevant aspects. (Example: If the possession of money suddenly is more of a danger than of an advantage (f.i., because all rich suddenly are labelled "thieves" and to be hanged), then adaptation to the new circumstances will consist in avoiding the possession of a more-than-average amount of money.)

Remark: Adaptation phenomena are evidently found in the biological as well as in the psychological and social realms.

2.4. Personality and Behavior: Two Description Levels

Repetition: Since organisms are conceived of as automata, their outputs are understood as *reactions* to inputs (not necessarily immediately preceding in time), determined according to specific internal schemes (the internal 'states') which derive output-variable values from the input-variable values.

We do not here assume any particular embodiment for the organisms. All reactions of an organism (such as the anatomical and physiological details of the reactions of a human player to the servings of his counterpart; or such as the knee reflex in humans) which are typical of only a certain type of physical embodiments are called *embodiment-specific* reactions of behavior. Organismic behavior which is not embodiment-specific is called *embodiment-invariant*.

All concrete physiological processes are embodiment-specific. Embodiment-invariant organismic behavior is therefore seen to be simply an abstraction of embodiment-specific organismic behavior.

We ignore here the exact and concrete physiological and anatomic details of, for instance, the articulation of words, or of catching a ball. It is only the behavior of *abstract* organisms with *any* kinds of natural or technical embodiments which we are interested in on the psychological and sociological description levels.

All *psychological processes* are embodiment-invariant, by definition: psychological processes are systems of physiological processes described at such a level of abstraction that they are embodiment-

invariant and thus also non-manifest.

(Comment: *Social processes*, in turn, will be explicated as systems of psychological processes, described at such a level of abstraction that they are personality-invariant.)

The input-output mapping of an organism which generates its embodiment-invariant behavior is called the *personality* of the organism.

We distinguish between three major subfunctions of an organism's personality: The *receptive function*, the *cognitive function*, and the *executive function*. We assume all three to be realized in a single neural network; we disregard actual anatomical changes which translate inpulses from the neural net into material changes in the environment. (We are here **not concerned with constructing** a *material* robot, namely.) The executive function consists thus only of firing appropriate effector neurons. The receptive function consists only in the response of receptor neurons to stimuli from the organism's environment. The states of the receptor neurons are together the input to the neural net, the states of the effector neurons are together the output of the net. The essential part of the input-output mapping is determined by the states of the inner neurons.

We call the total network also the *personality network*.

Within the personality network we need to distinguish a central region, which we call the *central nervous system* (CNS). This region is vaguely distinguished from the rest of the network in that it is not directed: While the general direction of

connections is from the receptors towards the effectors, the CNS-neurons are cross-connected into all directions with other CNS-neurons.

Only if we distinguish a CNS can we speak of *apperception*, thereby referring to the firing of receptors and to the consequent relaying, delaying, multiplying, or inhibiting effectors of inner neurons upon an impulse. These inner neurons must namely be located between the receptors and the CNS (no inner neurons are receptors).

The part of the net comprising these inner neurons plus the receptors is called *perceptor*.

In analogy, we call *effector* the part of the personality network comprising the inner elements between (in the sense of impulse flows) the CNS and the effector neurons plus these effector neurons.

2.5. Accomodation

We use the word "accomodation" to denote the class of processes which 'filter' the inputs to the apperceptor.

Feedback in the personality network can be negative or positive. If it is positive, it has a sensitizing, threshold-lowering effect; the network becomes thus more 'attentive' to subsequent signals. If the feedback is negative, it has an inhibitory, threshold-raising effect, preventing a greater number of subsequent signals from being transmitted.

Holst's Principle of Reafference (Holst/Mittelstædt 1950) applies to such inhibitory feedback. In man there seems to be a greater number of inhibitory than facilitatory feedback circuits (Gerard 1960).

2.5.1. Attention and Vigilance

Our introspective realization of a 'directing' of awareness is an awareness of the selective raising and lowering of normal firing thresholds in response to sensations (organism-exogenous sensory stimuli) or proprioceptions (organism-endogenous sensory stimuli). We shall call such directed awareness 'attention'. (Miller 1969 stipulates a different use: He uses "attention"to denote directed awareness.) Attention may occur without consciousness (Cheek 1959).

Cherry(1953, cited in Norman 1969) and Grey/Wedderburn (1960, also according to Norman 1969) have shown that attention (in our sense) often is supported by what we shall call 'vigilance', namely a

conscious filtering of inputs on the basis of semantic and pragmatic rules of concept formation. Basic-concept combinations which are recognized as being meaningful ('messages') are preferred to meaningless combinations, and meaningful combinations which have acute instrumental value (that is: which are cues that can facilitate the solution of a current task) are preferred to meaningful combinations which only have a potential instrumental value.

2.5.2. Conspicuity and Memorization

In order to enable the learning of new signal types and meanings, we shall make use of Helmar Frank's (1964, p 52-55) ideas regarding the 'surprise value' of a (stochastic) result. He defines the 'surprise value' of a message as the message's information content in proportion to the uncertainty it removes. This uncertainty is maximal when all 'possible' result messages are equally probable. It is computed as the average of the information contents of all 'possible' result messages. The 'surprise value' of a result message which occurs when its information content equals the average, is a 'neutral' value: messages with lower values are considered 'trivial' and higher values 'interesting', the more markedly, the greater the deviation from 1.

The 'conspicuity' of an event is computed upon recurrent situations of uncertainty (a sequence of die throws, for instance; this sequence may be of length 1); the measure is the sum of the 'surprise values' of the event's occurrences.

We can assume that one's awareness of an event is directly proportional to its conspicuity. Observe

that conspicuity can be self-defeating: as an event occurs more often, its probability changes and thus also the uncertainty of the next situation.

We can also assume that the probability that an event will be interpreted as a signal is proportional to its conspicuity. And we can finally suspect that the conspicuity of an implication depends upon the conspicuity of the involved events.

The computation of conspicuity values of stimuli and stimuli combinations is presumably simple enough to be realizable in simple nerve nets, and that consciousness in not a prerequisite.

It seems reasonable to let signal memorization (that is: the 'depth of imprinting' of an appercept of a signal in the memory) be a function of the conspicuity of that signal.

3. Learning

Learning is a phenomenon which we can observe only in cybernetic systems; and vice versa can we say: a system is a cybernetic system if and only if it is *adaptive* (that is: if it is able to *learn*.)

(Comment: We shall in this text not make any distinctions between adaptiveness and ability to learn although it is imaginable that a system is *able* to learn but for some reason still non-adaptive, especially if one compares internal with external states or responses. "Adaptation" and "adjustment" shall be used to indicate the existence of de-facto effects of adaptive behavior.)

We explicate "learning" vaguely as denoting a characteristic of system-environment interaction, where the environment manages to cause a change in the system in such a manner that its subsequent behavior is modified. However, although the environment-caused changes must involve modifications in the embodiments (f.i., in brain-cell states), must a distinction be made between adaptations on the one hand, and embodiment *damages* on the other, which also can cause a system to behave differently. (A wounded animal behaves differently from when it still was undamaged, but we usually do not wish to say that the animal now had *learned* to behave in the modified manner). We shall therefore limit the class of influences which we allow as *teaching* influences to such which induce modifications in the system's internal state via the system's specific input devices ("sense organs" or "sensors").

It is obvious that it is not easy to approximate an adequate definition of "learning". The difficulty

lies in the fact that "learning" commonly is used to denote phenomena which need to be described on different levels of detail.

Similar ambiguities have led and still lead to controversies regarding the relationships between different theories of learning.

F.H.George (1970, p 27) defines *learning* as "the process of adapting to changing circumstances and the recording of success and failure in the process. This means that we envisage a system which has the necessary flexibility to adapt to a change in the environment, although we would also wish to distinguish between adaptation which is built in (instinctive, innate) from that which is aquired as a result of transactions between the organism and environment. 'Learning' is, in fact, a term normally applied to the latter category..."

If one allows the use of such macrolevel concepts as 'success' and 'failure', then it is obviously still easier to distinguish learning processes from processes which we would not call learning.

The problem of definition is, of course, a pseudo-problem in the context of simulations as long as the simulation is ongoing, but when it comes to interpreting the results (perhaps even computer-assisted), then an adequate distinction is again of relevance. Let us therefore make a few additional comments on existing theories and simulation models of learning, before we concentrate upon the specification of a concrete model of learning.

3.1. The Learning-Theory Labyrinth

Studies of biological systems have resulted in a multitude of learning theories. Not all of these have been formulated in such a way that they were conceived of as complementary; instead they seemed to be contradictory to other theories. We believe that, for instance, the argumentations *for* stimulus-response theories ("S-R theories") and *against* cognitive theories of learning, and vice versa, often have been misunderstood as competition between theories on the *same* description level, while we would rather see them as an (often unnecessary) argumentation for the use of one description level at the expense of the other. Such argumentation is, evidently, quite in contrast to the approach to these matters which we propose in this book.

Another controvery seems to be the ground for more fruitful debates. Although we in this context shall not investigate psychological problems more thoroughly, it must be of interest to us which point of view is the most plausible and/or most strategic for us to adopt in the controversy about the *necessity of rewards* and the *sufficiency of contiguity*.

Let us indicate in short the three principal points of view in this matter, not only to point to experimental evidence concerning animals and humans (which can be found, f.i., in Hilgard and Bower 1966, and in M.H. Marx 1970), but also with the synthesis of an artificial learning system in mind.

(a) Pavlov's famous 'classical' conditioning experiments demonstrated that the mere co-occurrence (in other words: time-spatial proximity) of stimuli could induce associations of similar responses to

both stimuli.

(b) Thorndike's and Skinner's experiments with 'instrumental' conditioning were concerned with the effects of extraorganismic reinforcements.

(c) Each of these two types of approaches to learning via conditioning has been adopted in a vast number of subsequent and varied studies, the subtleties of which (in design and interpretation) need not concern us here. However, there have also been attempts at unification of the two standpoints into one. Hull, for instance, wrote 1937: "The Pavlovian conditioned reaction and the Thorndikian associative reaction are special cases of the operation of the same principles of learning"(Hull 1937, p 17). But Hull's hypothesis was merely that an S-R connection was strengthened if and only if it co-occurred with a reinforcing situational factor.

Seward (1970, p 93) states more decidedly a third standpoint:"instrumental conditioning involves the classical conditioning of a motivational process that supplies positive or negative feedback to the instrumental response".

Razran (1965, p 207-235) has recently proposed a hierarchy of levels of learning, with habituation, sensitization, and inhibition at the most elementary (perhaps phylogenetically most primitive and/or ontogenetically earliest) level, classical conditioning at the next, reinforcement conditioning at the third, and cognitive learning at higher levels.

Taking the still wide diversity of hypotheses concerning the intra-organismic effects of conditioning into due consideration, it seems recommendable

to select a combined contiguity/reinforcement model. A Thorndike/Hull approach is, in any case, a logical necessity in models of conscious and willed adaptation. We leave the question open (for the experts to decide after further development) whether or how reinforcement models are to be resolved into contiguity models. But let us comment that a reinforcement notion (although not in the sense of "reward" or "incentive" nor "punishment", but simply and more literally of "strengthening") also is logically y necessary in contiguity models. If an adaptive system, namely, is to make any distinction between available response alternatives, there must be a non-random criterium of choice between these alternatives. Conditioning, and also more complex forms of learning are processes by which response-choice criteria are somehow embedded in the system embodiment. (How this is exactly done in biological systems is neither undisputed scientific knowledge, nor of importance in this text.)

The model which we are led to choose is thus rather conventional: An adaptive system will somehow categorize environmental reactions to its behavior into three classes: Rewards, punishments, and neutral reactions. It will in the course of time tend to behave in such a manner as to maximize the frequency of rewards and to minimize the frequency of punishments; if necessary, it will compromise on a specific reward-punishment ratio other than the ideal 100%:0% ratio.

This means, in operational terms, that if a positive stimulus (a 'reward') is received as a reaction to the system's last output, then it causes an increase of the last output's probability of reoccurrence. A negative stimulus (a 'punishment')

as an environmental reaction in analogy, causes a decrease of the last output's probability of reoccurrence.

Contiguity plays obviously a strong role in reward models. We ask ourselves whether reward-learning can occur when environmental reactions are time-spatially remote (f.i., are delayed) but observable. Evidently,must our answer, from introspection, be "yes" for learning in general; experiments by Wolfe (1934), Riesen (1940), Perin (1943), Perkins (1947), and Grice (1948) indicate that relatively small temporal delays of rewarding stimuli do not inhibit reinforcement altogether, although delay length and reinforcement strength were found to covary reciprocally. But neither here are data conclusive; other experiments (f.i., Garcia/Kimeldorf/Hunt 1961), with a variety of animals, seem to indicate that delays may be considerable without eliminating reinforcements altogether.

It seems to us that ability to learn from time-spatially remote environmental reactions might require 'cognition' with capacities that are approximately in proportion to the extent of remoteness and inconspicuousness.

3.2 Three Simple Models of Elementary Learning

(a) A Model of Classical Conditioning

Let S be a fuzzy automaton called 'arc' which receives inputs called 'stimulus qualities' and which emits outputs called 'response qualities' (or simply 'responses'). The input set of S comprises at last 2 stimulus qualities. The operation rule of S is based upon a count of which response qualities S emitted in the past in reaction to each input. The output is sampled from the response-quality set with these past-response counts as sampling weights. The response quality which has been most frequent in the past will thus also be most probable in the future.

This is a trivial model, of course, but it shows very clearly the limitations of such pure contiguity conditioning. Since "reward" and "punishment" are unknown concepts in the universe of this model, there is no possibility to reverse conditionings which are stronger than all other reflexes.

In more elaborated ('de luxe') models of pure contiguity conditioning one would find memory-decay functions and perhaps non-linear accumulation of weights and a more sophisticated response-selection function.

(b) A Model of Instrumental Conditioning

The model of instrumental conditioning can be derived from S. Assume another fuzzy automaton, T, which has two input sets, namely one of 'stimulus qualities' and another of 'reinforcements'. T's operation rule in reaction to a stimulus quality is the same as that of S, but the reinforcements can induce changes in the 'bookkeeping' of past responses (so that even negative 'frequencies' are possible to occur).

For instance, let the 'book keeping' of past responses be stored in the form of lists such as "(x, (p, 14), (q, 9), (r, 11))".

In the case of this example, both S and T would tend to respond "p" to the input "x", and both would thus cange their internal state ('memory') by updating the above list into "(x, (p, 15), (g, 9), (r, 11))"

However, assume that reinforcements to T are numerical values between -1 and +1. Now, if the reinforcement upon the response p(given x) is w, the above list will in T be modified further into "(x, (p, 15+w), (q, 9), (r, 11))".

S coincides obviously with T if all reinforcements are zero.

(c) A Model of Response Differentiation

We can elaborate T still a little bit further, to obtain an automaton W which 'learns' new responses.

We also know that the reinforcements either are of the form which T accepts, or of the form "(z,w)", where z is a response name, and w a reinforcement value (as above). We call such reinforcement pairs 'instructions', because they are used to suggest that the last output should have been z.

Assume, for instance, that W behaved like T in the last example, but that we would have liked it to respond "z". After having received our instruction, W will store the list
"(x, (p, 15-w), (q, 9), (r, 11), (z, +w))".
Observe the reversed sign in w.

Even if w is +1 in these instructions, it will take many repetitions to override past response habits for good, however. A decay function (which substracts a small proportion of each count with each time step) could possibly avoid that the dominance of often-repeated reflexes becomes too strong. We shall also assume that no response selection occurs if the instruction is presented simultaneously with the stimulus-quality input. Learning will thus be somewhat faster.

Can W be made to 'learn' new responses *automatically*? This requires obviously that W 'feels a need' for a new response; that is: we must enable that W can be, 'dissatisfied' with its response repertoire and that it therefore increases it.

Let us therefore first make a trivial modification of W. Compute for each of the memory lists the sum of all counts in the list. For instance, the total count for "(x, (p, 2), (q, 4))" is 6. Compute then for each response the ratio of its count and the

total count; it will be 2/6 for p and 4/6 for q in our example. These ratios are the sampling weights of the responses, but since they may have been subject to reinforcements, they can also be interpreted as being 'adequacy measures'. We need now simply set an 'adequacy limit' for old responses, the value 0.5, for instance, and to stipulate that W 'invents' and emits a new response if no old response has an 'adequacy measure' that exceeds the limit. We assume further that a new response will be invented and emitted if the total count is zero.

In order to give new responses a chance, we must now store a 'specific count sum' for each response alternative. These 'specific counts sums' start from 1 when the alternative is added to the response set; *all* 'specific counts sums' are updated each time *any* response is emitted or reinforced.

For example, assume the empty response-disposition list "(x)" to be given from the start. A new response will be invented and emitted because the total count is zero. We obtain thus "(x, (r1, 1, 2))". The "2" is r1's 'specific count sum'.If no positive reinforcement occurs, W will generate the list "(x, (r1, 1, 3), (r2, 1, 2))" next time, et cetera, until a new response receives a positive reinforcement.The older responses can be made to decay away and to reoccur as new now and then.

We must, of course, assume that W is associated with a limited response repertoire from which the new response can be picked at random.This repertoire must be thought of as being somehow defined by embodimental constraints.But since it is not appropriate to speculate into the details of this de-facto definition in a psychological-level study we shall abstain from doing so.

4.Pattern-Recognition

Pattern recognition ("PR") is a topic which could be treated under the heading "perception" as well as of "learning"; PR-models usually simulate both phenomena, most often behaviorally, not structurally. As we shall see are PR-models very well suited as models of the elementary forms of learning which we discussed recently; some models point simultaneously towards possibilities of synthesizing cognitive processes from elementary (conditioning) ones. It is therefore opportune to worry little about structural inadequacies of PR-models as models of elementary human learning.

Structural similarity is, of course, an ultimate aim, but in a default model we can already be satisfied if we achieve behavioral similarities between the mental processes in humans and the simulating computations in the model. Once we do have a complete and in its psychological-level parts behaviorally adequate model, it is very easy to replace such parts with structurally more assimilated part-models of human mental processes. While constructing a first model it must be of highest priority to obtain *coherence* (meaning that all important model parts exist and fit together), even if this order of priorities should result in a model which has only behavioral adequacy (at the psychological description level). The last three decades of psychological and sociological research have shown that it is easier to elaborate part models than to achieve coherence in a complex model. We expect that, once we have a coherent multi-levelled model, we shall never lose coherence in our total model, no matter how far we elaborate and modify the very first version.

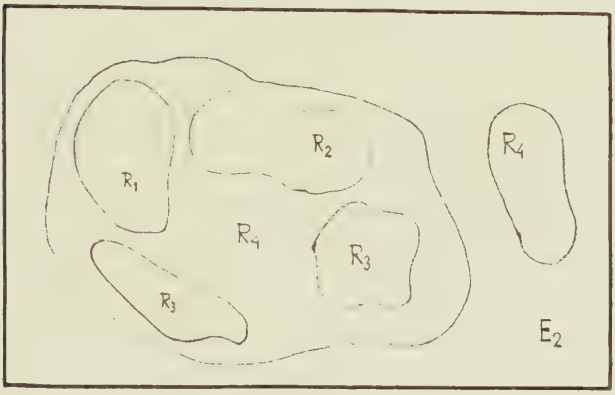
4.1 Trainable Pattern Classifiers

We shall assume that each set of data to be classified is an ordered set of n real numbers x1, x2,...., xn. Such a set is called an n-dimensional *pattern*; the set members are called *pattern components*. Any device for sorting patterns into categories will be called a *pattern classifier*. We shall label the categories by the integers 1, 2,...., r. "0" may denote a null category into which the classifier sorts patterns which, according to the furnished criteria, are not unambiguously classifiable.

Any n-dimensional pattern can be represented by a point in an n-dimensional Euclidian space E^n. E^n is called the *pattern space*. (Equivalently, the *vector* extending from the origin to the point (x1, x2,...., xn) can also be used to represent the pattern. We shall denote both the pattern point and the pattern vector by the same symbol).

A pattern classifier is thus a device which codes the points of the pattern space into the category numbers 1, 2,...., r. Let R_i denote the set of points in E^n which are mapped into the number i. Each category number is hence associated with a set of points (called *regions*) in E^n, denoted by one of the symbols R_1, R_2,...., R_r. (Mappings where r is infinite, or where not most of the regions contain more than 1 point, (that is, where any slight motion of a point changes the category of the point) cannot be treated appropriately in these terms and are not considered by us here). We assume that the boundary surfaces of the regions containing points of the same categories belong to both regions.

The regions on E^n may be disconnected:



The boundary surfaces (i.e., Jordan curves in E^2) separating different regions from each other we shall call *decision surfaces*; the regions into which E^n is divided by these decision surfaces we shall call *decision regions*. (from: Nilsson 1965, p 4-5.)

4.2 Apperception

We conceive of apperception as a process of stimulus-pattern classification (that is, as an abstraction), the result of which is a concept. It is presumably wrong to believe that apperception is a single-level process; we believe it to be very plausible that apperceptors are ubiquitous in all afferent channels (and also in the efferent ones,as we shall show later). Anyway, we shall stipulate that an 'appercept' is the *final* result of apperception; the 'appercept' is what is transmitted to the CNS. The intermediate apperceptor outputs will not be called 'appercepts'.

We indicated earlier that appercepts can be assumed to be the elements of reasoning, the basic concepts. It is probable however, that not all basic concepts have names; we know from the reports of people (like Helen Keller) who learned language late in their lives that concepts do not need to have names to be usable.

We wish to point out that our conceptual system is perhaps not quite satisfactory. Concept formation can namely also be modelled in terms of the same pattern classifiers with which we simulate apperception. The only reason for calling some of the abstraction 'apperception steps' and some of them 'concept formation' can be then that some of the involved categories have names.

We shall postpone the further discussion of concept formation until the next chapter.

4.3. A Model of an Apperceptor

An apperceptor is an automaton composed of elementary learning modules called 'sensors' and which are similar to the 'arcs' we have described earlier. Sensors differ from type-W arcs (cf. section 3.2, c) in that they are not autonomous and that the responses referred to in reinforcements instructions and in disposition lists therefore are apperceptor responses, not sensor responses. The sensors' outputs are instead the relevant disposition-lists. For instance, let a particular sensor's 'memory' contain (A, (c, 4), (a, 2)), (B, (e, 3)) where "A" and "B" are sense qualities, c, a, and d are response qualities which the apperceptor has manifested in the past, and 4, 2, and 3 are frequencies of occurrence. The sensor's output upon the input "A" will be "(A, (c, 4), (a, 2))", upon the input "B" it will be "(B, (e, 3))".

We disregard here the problem of automatic respons response learning (in the manner of W-type arcs), merely to make our model's presentation as simple as possible. The automatic response learning ability is easily added to the apperceptor as a whole later on, since its final choice of response is based upon a sort of 'adequacy list'.

The operation rules of apperceptors are as follows: Given be an input vector, wherein each component is a 'sense quality' which is the input to one of the sensors. There is a one-one correspondence between the input-vector dimensions and the sensors. Each dimension can have its own specific range of values. Reinforcements and instructions submitted to the apperceptor are

simultaneously valid for all sensors alike.

The sensors respond to sense-quality inputs by outputting lists, we said. The apperceptor scans all these lists and each mentioned response's absolute frequency is transformed into a relative one, by dividing it with the total number of past trainings of that response times the numbers of sensors and these relative frequencies are accumulated separately for each response. The apperceptor's output is eventually the response with the maximal accumulated relative frequency.

4.3.1 Example: Unambiguos Training

Submit the following instructions to the apperceptor:

((a,d,g),(C,1)), ((a,d,k),(F,1)), ((a,b,g),(A,1)),

((a,b,k),(G,1)), ((a,d,k),(F,1)), ((f,b,g),(D,1)),

((f,d,g),(H,1)), ((f,b,k),(E,1)), ((f,d,k),(B,1)),

((a,d,g),(C,1)), ((a,d,k),(F,1)).

The apperceptor consists evidently of three sensors. Its output repertoire comprises the 8 response qualities A,B,...,H. The "1" indicates that the instructions are given reinforcement 1.

The memory content develops as follows:
H is after the first instruction:

(((a, (C,1))), ((d, (C,1))), ((g, (C,1)))

Total numbers of trainings:

((C,1))

After the second instruction:

((a, (C,1), (F,1)))

((d, (C,1), (F,1)))

((g, (C,1)), (k, (F,1)))

Total numbers of trainings:

((C,1), (F,1))

After the third instruction:

((a, (C,1), (F,1), (A,1)))

((d, (C,1), (F,1)), (b, (A,1)))

((g, (C,1), (A,1)), (k, (F,1)))

Total numbers of trainings:

((C,1), (F,1), (A,1))

Et cetera. After the tenth instruction:

((a, (C,2), (F,3), (A,1), (G,1)), (f, (D,1),
(H,1), (E,1), (B,1)))

((d, (C,2), (F,3), (H,1), (B,1)), (b, (A,1),
(G,1), (D,1), (E,1)))

((g, (C,2), (A,1), (D,1), (H,1), (k, (F,3),
(G,1), (E,1), (B,1)))

Total numbers of trainings:

((C,2), (F,3), (A,1), (G,1), (D,1), (E,1), (B,1)).

Assume that we now submit the input "(a,b,k)" to the apperceptor. Since this is no instruction, the sensors will produce outputs, namely the lists

(a, (C,2), (F,3), (A,1), (G,1)),

(b, (A,1), (G,1), (D,1), (E,1)), and

(k, (F,3), (G,1), (E,1), (B,1)).

The apperceptor will transform the absolute frequencies into relative ones, on the basis of the total numbers of responses, multiplied by 3 (namely the number of sensors):

(a, (C,2/6), (F,3/9), (A,1/3), (G,1/3)),

(b, (A,1/3), (G,1/3), (D,1/3), (E,1/3)), and

(k, (F,3/9), (G,1/3), (E,1/3), (B,1/3)).

Accumulation leads to the list

((C,2/6), (F, 6/9), (A,2/3), (G,3/3), (D,1/3),

(E,2/3), (B,1/3)).

The response "G" is found to have the maximal weight and is therefore made the apperceptor's output.

4.3.2 Example: Ambiguous Training

There is nothing remarkable about an apperceptor that stores pairs and responds with one half if fed with the other. Already the simplest arc can do that. But we shall now show that an apperceptor really is a pattern classifier: given a set of possible responses, it selects the one for which most clues speak.

Thus, if we had trained the apperceptor with a sequence of instructions which all had random errors, we should still obtain the same correct responses from the apperceptor.

For instance, submit the sequence of instruction that should have been the ones we gave before but which somehow have become distorted during transmissions (we omit commas and parentheses for readability; all reinforcements have strength 1):

fbg:A, adk:C, adg:A, abk:G, abk:A, fdg:C, adg:B,

fdk:H, abk:E, fbg:G, abg:C, fbk:G, adg:H, adk:B,

fdk:E, fbg:H, adk:G, fdg:B, fbg:E, fdg:B, fbk:B,

abg:G, fbg:H.

The apperceptor's 'memory' content will be after this sequence:

((a, (C,2), (A,2), (G,3), (B,2), (E,1), (H,1)),
(f, (A,1), (C,1), (H,3), (G,2), (E,2), (B,3)))
((d, (C,2), (A,1), (B,4), (H,2), (E,1), (G,1)),
(b, (A,2), (G,4), (E,2), (C,1), (H,2), (B,1)))
((g, (A,2), (C,2), (B,3), (G,2), (H,3), (E,1)),
(k, (C,1), (G,3), (A,1), (H,1), (E,2), (B,2)))

The total numbers of training are:

((A,3), (C,3), (G,5), (B,5), (H,4), (E,3)).

Response qualities D and F have not been taught yet, and the apperceptor will therefore respond "C" to "adg" as well as to "adk", and it will respond "H" to "fdg" as well as to "fbg".

For instance, "adk" will make the sensors reply with the lists for a, d, and k, which will be transformed and accumulated into

((C,5/9), (A,4/9), (G,7/15), (B,8/15), (E,4/9),
(H,4/12))

The apperceptor's response will evidently be "C", but "B" (correct reply to "fdk") and "G" (correct reply to "abk") are seen to be very close rivals.

We wish to point out that such behavior is theoretically interesting for the interpretation of educational influences in general. One wonders for instance, how children learn sex roles. We suggest that sex roles are 'pattern categories' that are derived from very many incomplete and

sometimes even contradictory examples.

4.3.3 Example: Inverted Use

Apperceptors can be used in an inversed manner to create 'ideal types'; the inverse apperceptor's response to a response quality as input is a 'typical' sense-quality input vector: Thus, if we give the input "G" to the inverse of the apperceptor that has been trained as was illustrated in the last example, we obtain the output "abk".

The operation rules for the inverse of an apperceptor are very simple: in the apperceptor's memory content, look up the weights of "G" in the first sensor's all lists. In our case we will obtain merely two values because the sensor has, as yet, only received two different sense qualities, namely "a" and "f". "G" has the weight 3 in the list for "a" and the weight 2 in the list for "f". The inverse apperceptor selects the value for which "G" has the higher weight, namely "a", and procedes to the next sensor. It yields, in similar fashion, the value "b". The third sensor yields "k".

4.4 Comments

Apperceptors of the described kind are called 'linear pattern classifier' (by Nilsson 1965, p16) or 'Learning Matrix' (by Steinbuch 1963, chapter 14). Steinbuch has also shown that 'Learning Matrices' can be coupled in various ways (Steinbuch 1963, p226, p229, p234).

The outputs of several apperceptors can be inputted to another apperceptor, and we can thus achieve abstraction in steps. Or we can achieve a kind of 'translation' by coupling an inverse apperceptor to a non-inversed one, or a non-inversed one to an inverse one.

It happens that parts of input patterns to apperceptors are missing during training or during use. When this is the case, the concerned sensor is simply ignored. This allows us also to teach patterns of varying lenghts.

4.4.1 Inductors and Deductors

It would be confusing to speak of 'apperceptors' and 'inverse apperceptors' in contexts outside the perceptor. We shall therefore stipulate that apperceptors are *inductors* that belong to perceptual channels. Inductors are all to have exactly the operational definition that we specified for apperceptors. We introduce the terminological differentiation only because "apperceptor" is shorter than "inductor in the perceptor". An inversely used inductor will consequently be called '*deductor*'.

4.4.2. Concepts and Conceptions

A concept is (in epistemological contexts) usually regarded as a class of instances or qualities (that is: a set of values). The 'conception' associated with a concept is identified as being the 'content' (also called 'intention') of that class. Is it possible to reflect these epistemological notions in our psychological context?

We understand *induction* to be the transformation of a (possibly not complete) extensional definition of a class of phenomen into an intensional definition. This can be done by presenting a sequence of class members to an inductor which will derive a concept from it. The conception of this concept, obtained by using the same inductor inversely as a deductor, is the semantic correspondence to an intensional definition of the formed concept.

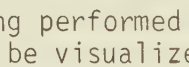
In order to allow for *fuzzy* and ambivalent conceptions of a concept in an individual, it seems appropriate to say that the conception of a concept is a fuzzy set (cf. Volume 2, chapter 2.) sampled from the list of possible deductor responses to the submitted concept. If the deductor produces several 'ideal types' that have very similar weights, the conception will obviously not be very clear.

The *verbalization* of a conception can be more or less trivial, depending upon whether the involved concepts are named. Unnamed concepts must be resolved further into conceptions, and if these heuristic searches fail, alternative conception elements must be tried.

5. The Personality Network

In organisms there is constantly non-manifest behavior (=internal-state transitions) going on. Our discussion of pattern-categorizing systems indicated how complicated internal states could be abstracted. We also assumed the existence of a central part of the personality network "CNS" which expedes fast intra-CNS communication in terms of abstracted state-descriptions.

Let " " denote an inductor and let

"  " denote a deductor. The information-processing performed in the personality-network can then be visualized as is done in figure 5.0.

That is: the organism receives, filters, and abstracts exogenous stimuli (in the perceptor), processes the abstractions in the CNS, which eventually triggers 'appropriate' responses, i.e., signals which propagate in the effector. We have in the figure below for the sake of transparency omitted channels for the processing of endogenous stimuli; these channels may *also* contain apperceptors and they always connect (and influence) apperceptors belonging to afferent or efferent channels for exogenous signals.

The motoric effector consists mainly of deductors. These must be trainable similarly to inductors. The patterns emitted by the effector are patterns of motoric responses.

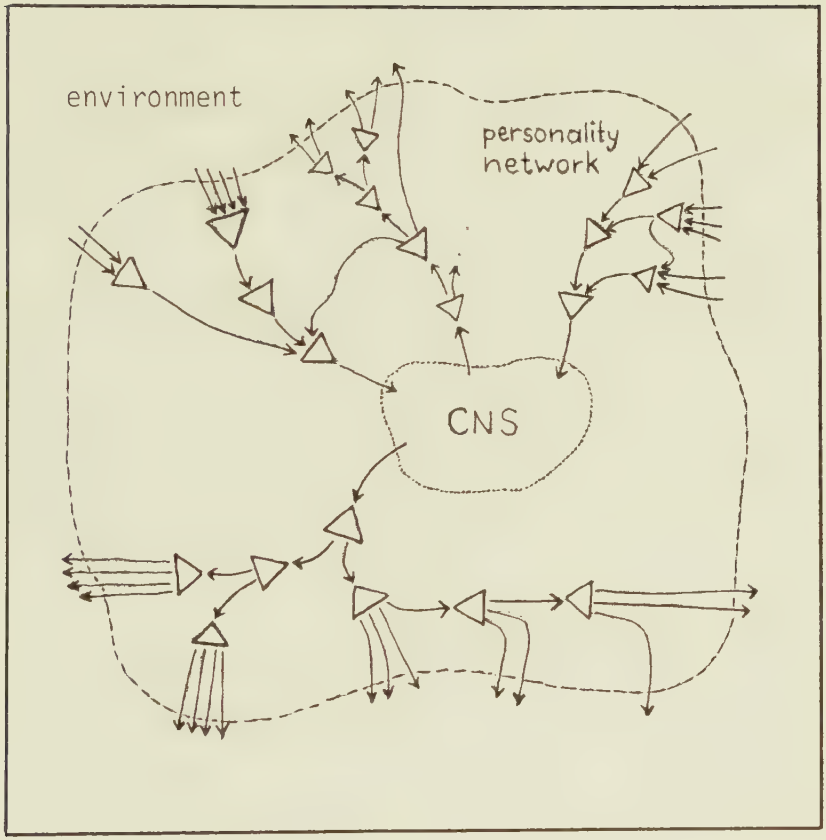


figure 5.0.

5.1 Thinking and Reasoning

Thinking is non-manifest behavior; "*thinking*" is a collective term used to denote intra-CNS communication processes. A *train of thought* is a particular non-empty sequence of intra-CNS communicative steps. (If one wishes one can distinguish further f.i., between *deep* thinking (involving long trains of thought) and *shallow* thinking (involving only a few trains of thought) and *broad* thinking (involving a multitude of train trains of thought); between *consequent* thinking (involving trains of thought with few loops) and *inconsequent* thinking (involving trains of thought with many loops and loops of loops).)

We do not wish to stress the traditional division of 'thinking' into 'reasoning' and 'imagination', because we believe that reasoning is very clumsy as long as the imaginative ability is lacking which provides us with instantaneous 'maps' of the context we speak about. It is not improbable that inversed use of apperceptors is involved in imagination generations; it seems namely that the Gestaltists' law of 'good form' of visualized objects obtains a very satisfactory explanation from the behavior of inverse apperceptors.

5.2 Problem Solving and Planning

Problem solving is, generally speaking, a problem of transforming an initial pattern or state into a goal pattern or state. This sounds misleadingly easy. (So does also the formal problem-solving theory proposed by Mesarović (1965).) The difficult part of problem solving is namely the *identification* of 'appropriate' methods (transformations) to be used in given situations, both initially and on the way towards the solution.

This involves the subproblem of identifying the given situations. Apperceptors can easily be used for giving a 'robot' the capability to select appropriate methods in one-step problems. Situational, cultural, and subcultural particularities can adequately be taken care of. Problems with very few steps can be solved by the trial-and-error method. But are there generally effective problem-solving strategies available for the 'robots'?

The last two or three years have seen the implementation of several effective 'problem-solver' designs on the basis of theorem proving (cf. Volume 2, section 12.). Such a 'problem-solver' is given a goal description in predicate-calculus notation and this formal description is treated by the 'problem-solving' program as a *theorem* which is to be proven true or false. The proof is to be based upon the present-state description contained in a certain database and also upon a given list of expression-transformation rules ('operators').

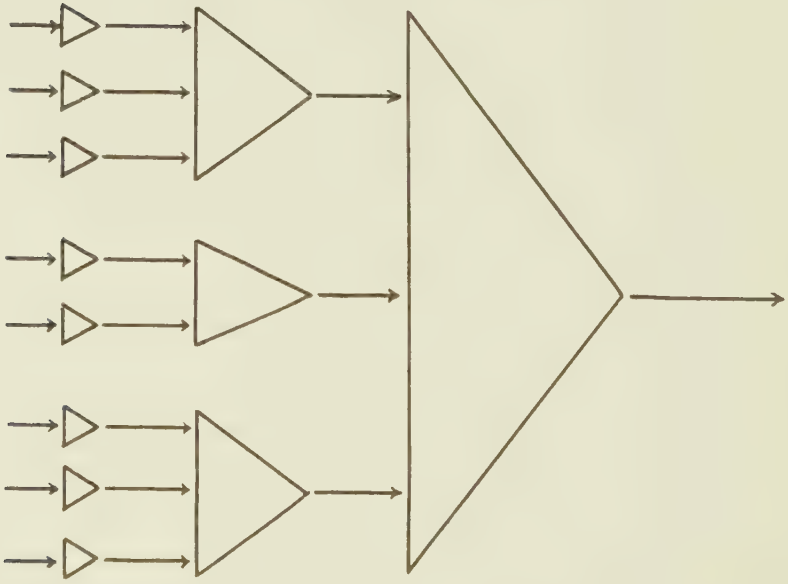
Models of human behavior require such perfectly reliable problem solving only for being able to simulate explicitly rational problem-solving behavior (the elaboration of an important administrative decision by some 'robot', for instance). But in the normal course of simulation, there is no need for scientific certainity of the feasibility of a plan. People do not first *prove* that they are *able* to reach a certain shop, each time they intend to buy something there; but they do, of course, assume that they can because they usually can and because nothing extraordinary has happened which can prevent them from getting there. In other words: People will start to implement their intentions if they fail to be able to prove, or are not interested in proving, that there are obstacles which will impede them from reaching their goals. This is the way we intend to tackle the planning problem.

5.3. Conscious and Subconscious Thinking

We believe that humans usually solve problems on two levels simultaneously: on a *conscious* level, where named concepts and relations are employed for problem representation, and a *subconscious* level where unnamed concepts are processed by cascaded inductors (some of which *may* output named concepts, however) and deductors. Conscious problem-colving involves the manipulation of a semantic database and leads to a focusing upon certain beliefs (that is: conceptual structures in the database). Subconscious problem-solving may lead to the triggering of 'feelings', 'memories', and acute behavior dispositions. (Rohracher (according to Eysenck/Meili 1972, keyword "subconscious") seems to support such a view: he calls the subconscious 'co-consciousness'.) That is why we can rationally believe something to be correct and in order, and yet have a 'funny feeling' about it. Subconscious problem solving is fast because relationships between concept components are not differentiated; relationship types can therefore not be reasoned about subconsciously (if our beliefs about the subconscious are adequate). Subconscious thinking should therefore be minimal when we are occupied with abstract relational problems. It is probably undirected, or at least not strictly directed, and it is therefore not as *narrow* as conscious thinking. We can suspect, however, that the conscious direction of awareness ('focusing') upon certain structures within the database vaguely influences the subconscious processes, which thus maintain a less differentiated, diffuse 'contextual halo' of 'associations' around the consciously focused beliefs.

5.3.1. Concept Formation

We believe that subconscious and conscious thinking involve the very same database and conceptual structures, but that they differ in the manner of using them. That implies that there is only a single type of concept-formation process. We assume that concepts are formed in cascaded inductors, as illustrated in figure 5.3.1a:



A crucial question is how relations are induced. Are they abstractions from set extensions? There is, of course, always a basic relationship between linked concepts, namely that of being linked. This relationship can be differentiated by all kinds of imponderable but relation-specific cues, if these are added to the pattern that is to be abstracted into a new concept. The interesting thing is that we evidently can distinguish the components of the 'higher' concepts, although relations always appear to us to be less tangible than the related arguments. Also the subconscious can presumably identify the components of a concept.

5.4. A Simple S-O-R Model

We describe the following simple model mainly for the purpose of illustration of the type of 'holistic' processing which we believe to be typical of subconscious thinking, in contrast to the 'atomistic' analyses of conceptual structures which characterizes conscious thinking. The outlined model can be said to be of the 'stimulus-organism-response' variety. It suggests also how we wish to let motivation be a basis or context for cognitive behavior.

Assume the following standardized but general problem-situation: An individual believes himself to be in a situation of type r. He is at present characterized by a need-vector (or utility-vector or goal-vector) U with n components; we may assume them to correspond to the (not mutually orthogonal) dimensions

(1) physical utility
(2) physiological utility
(3) psychological utility
(4) micro-social utility
(5) macro-social utility
(6) economical utility
(7) political utility

The individual uses this utility-vector for the evaluation of alternative concrete, mutually inconsistent goals or subgoals. The individual does not necessarily have an ultimate concrete goal in his life, but *if* he has, other goals can be regarded as subgoals of this goal. We must not be so unrealistic as to assume that all, or most, of this individual's goals necessarily are subgoals of the concrete ultimate goal; it is therefore appropriate to associate each goal with an aspiration-intensity estim-

ate, which specifies the strictness with which that goal is persecuted. If the individual does *not* have one concrete ultimate goal in life, he may have a *set* (containing at least two elements) of concrete long-term goals; these may in practice be mutually exclusive but may still be aspired simultaneously or by turns. The individual may even lack such long-term goals and only have situation-induced concrete short-term goals.

From where does the individual obtain his goals? It is obvious from earlier chapters that biological needs are dual to 'abstract' goals; concrete goals can satisfy biological needs to varying extents, and the individual tries to implement his 'abstract' goals optimally in terms of concrete goals. Conditionings, on various levels, can give rise to 'psychological needs', f.i., ethical or esthetical ones; curiosity may also be such a 'psychological need'. Responsibilities, norms, and obligations of all kinds, including constraints which the individual has not internalized, can be regarded as 'sociological needs'. The evaluation of concrete goals must be the result of an *internal process of bargaining* between the subsystems generating these different 'abstract' goals. This evaluation procedure can be operationalized in terms of inductors as indicated by the figure 5.4 on the next page.

The goal alternative s is simultaneously submitted to n (here 7) inductors, one for each of the above-mentioned utility dimensions. The inductors must by then have learned the approximate utilities of a great number of states. The resulting utility-vector is an over-time estimate which is weighed with the *present* need-vector U and then given as input to an (n+1)st inductor which has been trained

to combine the n aspect-specific evaluations into a single one (which we shall call *desirability estimate*.)

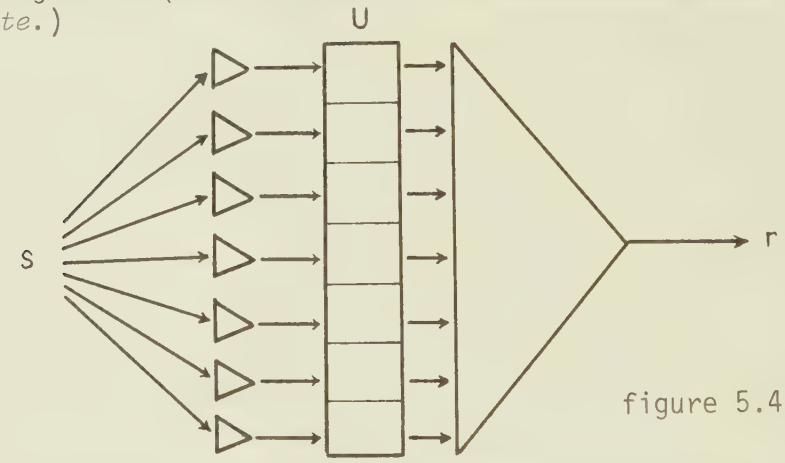


figure 5.4

This does still not answer the question how the individual obtains the sets of concrete goal alternatives. He definitely does not sequentially test all possible states for their desirabilities. Nor does he merely have a list of states which he has *heard* to be goals for others. But it has presumably something to do with his earlier experiences, immediate or heard-of. It is therefore not far-fetched to suppose that the individual uses the left-hand inductors in figure 5.4 in reverse mode, in order to learn about idealized states of high and low utilities. We shall call such idealized states of high utilities *concrete ideals*; concrete ideals represent half an operationalization-step from 'abstract' to concrete goals. Concrete ideals need not be realistic; we call a concrete ideal *subjectively realistic* if the individual knows *that* there is a possibility to get from the individual's presently given situation to the better circumstances of the regarded concrete ideal; the individual need not

know *how* to attain this concrete ideal.

The individual's knowledge of the fact *that* it is possible to transform one state into another can be recorded in a transition matrix. All kinds of manipulations with such transition matrices must then be made available to the simulated problem solver. The individual has thus a chance to estimate *how* realistic (subjectively, of course) a not directly realizable ideal is: repeated transition-matrix multiplications will tell him so.

6. Semantic Databases (Belief Systems)

We shall now concentrate on conscious thinking ('reasoning'). Reasoning is 'atomistic' in the sense that it does not 'look at' a whole structure at once, but scans over parts of it. These parts are 'reified', namely as named concepts and relationships between concepts. We know little about the exact differences between 'holistic' and 'atomistic' thinking but we know from introspection that there is such a difference. Distinctions such as between 'intuitive' and 'rational' thinking seem to parallel ours. Since 'atomistic' thinking is 'focused', there is ground for a suspicion that the difference between conscious and subconscious thinking lies in the degree of parallelity (as opposed to sequentiality) of thinking. But we do not know how such shifts could be implemented in the human brain.

We shall speak of 'beliefs' instead of 'concepts' when referring to certain conceptual structures whenever we want to indicate that we look at them in an 'atomistic' manner.

Semantic databases can best be visualized as graphs, where the nodes represent spoken-about phenomena, while the links are directed and labelled and represent relationships between these phenomena.(Unary predicates are represented as membership relationships between a phenomenon and predicate categories.) How these databases are materially stored is beyond our present interest because the implementation of semantic databases is a relatively trivial problem, while the specification of appropriate operation rules for them is still very problematic.

A semantic database can also be called 'belief system', which indicates that we look at it pragmatically, namely as implying certain truth values. This pragmatic aspect (at least what regards factual rather than logical truth) is often disregarded in purely linguistic database uses where one derives sentences from the stored conceptual ('deep') structures or vice versa.(Cf. Siklóssy 1972; Schanck et.al. 1973; Wilks 1973; Colby 1971.) Leech 1969 is a linguistic study, but the author examines also the logical relationships between semantic units, p 34-40 and mentions the extra-semantic contextual influence upon factual truth, p 85ff. Cf. also Sandewall 1972, p 1.6)

Consequently, we shall also use the terms "belief" and "semantic description" synonymously.

6.1. Semantic Descriptions

Much effort is nowadays put into research on linguistic methods of assessing the meanings of natural-language sentences. It is understood that information systems cannot store source texts in their original forms if the systems are to work effectively on the stored information in order to draw (correct) conclusions. Since meanings can be expressed in very many different 'surface-level' forms, one expects to be able to avoid multiple storages of a message that has been formulated in several different ways. One expects furthermore that the specification of inference rules will be essentially easier if the variation in form is strongly reduced. One tries even to specify a complete 'vocabulary' of semantic elements which are to be building blocks of any meaning or belief. Schank et. al. (1973), for instance, postulate only six basic concept types: things, actions, attributes of things, attributes of actions, times, and locations, and these concept types can be connected in only 16 possible ways. Wilks (1973) proposes 60 semantic units. Leech(1969) has given a very thorough and comprehensive semantic description of natural English and suggests a great but still manageable number of descriptors. (We believe that Leech's general approach and categories are the as yet best choice for our own database's representations.)

Also Lakoff (1972) has studied the logical forms that determine the meanings of natural-language sentences.

Lamb (1970) has proposed a remarkably easy-to-realize 'cognitive-network' design which is based upon class inclusions and which promises to render

it very easy to make inferences (cf. figures 22, 23, 24). Lamb's networks may not be able to catch subtleties, but they suggest that an effective cognition implementation perhaps will comprise an interplay of several, differently structured, networks.

As an example of the linguistic use of semantic representations, let us mention Siklóssy (1972), who describes an automatic translator which translates semantic tree structures into natural-language sentences. This translator is also very easily trainable to express itself in new languages.

Very recently (that is, in March '74) we have also happened to obtain a copy of a paper on novel-writing by computer (Klein et. al. 1973) which in fact describes a system where a monitor makes modifying inferences in a semantic database and then generates natural-English descriptions (murder stories) of the invoked changes. This system is thus in its intent very similar to our own atman system, the two are, in fact, largely complementarily, it seems.

6.1.1. Context and Meaning

Semantic descriptions are no longer context dependent; dependency upon the context needs only be observed when one derives the semantic meaning of a sentence. The context can guide the translation but all referred-to contextual information will eventually have become embedded in the resultant conceptual structure.

6.2. Semantic Inferences

Sandewall (1972) is, to our present knowledge, the only author who has specified a comprehensive set of rules for semantic inference. Findler/Chen (1971) specialize on temporal relationships and causality. Leech (1969, p 34-43) mentions a few general rules for testing consistency and making inferences in a database structured according to his specifications.

More research is necessary on spelling out general semantic-inference rules for regulating inference processes in semantic databases. But starting with Sandewall's collection of general rules, one can incrementally add rules that are specific for one's own field of interest at least.

6.3. Pragmatic Inferences

It is possible to let an incomplete simulation model become the content of a semantic database. Instead of programming all its subroutines, we could then more simply describe the intended behavior in natural language and make extensive use of the system's capability to perform semantic inferences. This requires the implementation of a 'command mode' in the system which enables us to modify the database in the course of inferences in such a way that it reflects a different factual state than it did before.

The brick-world system by Terry Winograd (1972) is an implemented example of a simulation system that can be 'programmed' by giving it commands in natural English, and which also can reason about its own 'actions' and ask questions or give answers in correct English.

Winograd's brick world resembles the world of the STRIPS robot (described in Fikes/Nilsson 1971; cf. also Coles 1969), and both approaches to robot intelligence are also very similar in their principles. Both achieve problem solving by theorem proving; or, to be exact, their robots are said to obtain a *plan* of the route to the material-world problem's solution when they find a proof of a theorem that describes the material-world's goal state. The very problem remains thus unsolved until the plan has been executed in its entirety.

6.3.1. Implicational Molecules

Abelson/Reich (1969) have directed our awareness upon the existence of what they called 'implicational molecules'. The ascription of motives for action (studied by Jones/Davis 1965 and Steiner 1970, for instance) is one most important example of the use of implicational modules. Very recently have Schmidt/Addamio (1973) in fact proposed an attractive model of intention attributions.

6.3.2. Abilities and Requirements

We disagree with McCarthy/Hayes (1969) and others that have come to follow them (cf. the inputted declaration of "Pushable (box1)" in Fikes/Nilsson 1971, table 1) on the appropriateness of problem-solver inputs which explicitly indicate that certain things *can* be done. Such indications are camoflaged commands.

It would be much more adequate to assume that such indications are to be the result of reasonings ultimately based upon own experience and observations made by the problem-solving robots. Such indications are thus to be regarded as intermediate results of learning processes.

Freeman/Newell (1971) also use capability indications in their inputs, but not for telling a problem solver what *it* can do but rather what certain instruments in its environment can do. They also specify requirements for the realizability of certain instruments ('operators'). Given such information, a robot can itself draw the conclusions whether it

can do certain things or whether it cannot. It can also evidently *invent* new instruments this way, and possibly also new uses for old instruments.

7. Emotion

We regard emotions as patterns of subconscious reactions upon sensory inputs. (We ignore as yet the possibility that there might be a continuous and basic tonus in each individual.) These patterns are the outputs of deductors which have received cognitive inputs (similar in content: Bindra 1970, p7).

Most writers in the field agree nowadays that the labelled emotions we usually think of when we speak of emotions (cf. Davitz 1969, which contains a 'dictionary of emotions') are complex states; some suggest that some of these emotions are more basic than others. Robert Plutchik has proposed that emotions are mixtures of the eight basic emotional reactions *anticipation*, *anger*, *joy*, *acceptance*, *surprise*, *fear*, *sorrow*, and *disgust* (according to Murray 1964, p 56-57).

Caroll Izard proposes similarly that the basic constituents of all emotional patterns are the basic emotions interest, joy, surprise, distress, anger, disgust, contempt, shame/shyness/guilt, and fear (1972, p 2). Each of these is assumed to add a special emotional 'flavor' to specific life experiences.

These eight or nine emotions are certainly not 'basic' in the sense that they can not be resolved further into patterns by deductors. But we ignore whether the notion 'emotion' is to comprise the primary resolution (through a tree of deductors) of a cognition into visceral-somatic reactions as well as the feedback of these reactions into the CNS, or only the former or only the latter. This is namely merely a question of choosing an appropriate

convention. To speak on a more abstract level: We support the idea that emotions are vague classes of embodiment states that have been cognitively or reflexively triggered and which then are cognitively interpreted (cf. Bindra 1970, p8).

The findings of Bridges (1932) regarding the differentiation of affect in children suggest that children *learn* emotions.

Do emotions have any importance in robot psychology? We believe that they must have, at least in all cases where robots are used to simulate social interactions. Emotions can be regarded as spontaneous but vague subconscious evaluations of circumstances, and can thus be used in learning preferences.

8. Volition

"Will" (synonymously:"volition" or "conation") is presumably one of the still vaguest concepts in psychology (cf., for instance, the peculiar article on "will" by Mierke in the Eysenck/Meili/Arnold (1972) encyclopedia). When it is at all dealt with, it is done with little originality and only superficially (cf. Keiler 1970, which is supposed to be an innovating monograph on the subject; but the explication of "will" is allotted merely two of the (otherwise thorough) monograph's 208 pages; the very ambitious work by Miller (1967) does not even mention "volition" or "will".)

Hofstätter (1957, p 329ff) points out that the conceptions of 'volition' or 'will' differ between cultures. Certain positively evaluated things may be said to be 'willed', while negatively evaluated things can only be aspired on the basis of 'low motives' ('drives' or 'passions'). Dorsch (1959, p 328-329), on the other hand, distinguishes between 'Triebhandlungen' that spring directly from a single motive, and 'Wahlhandlungen' that are consequences of a choice among several motives. Also Hofstätter mentions that 'will' is usually said to be involved only in the cases of behavior choice.

But what *is* will? Is it a psychic process of its own, or is it the experiencing of certain patterns of other phenomena? Keiler and Hofstätter inform that both hypotheses have been proposed, but do not tell whether either one has been shown to be more valid than the other.

We are thus again out on our own if we wish to model volition. We suggest that 'will' is not to be regarded as a phenomenon of its own, but that it is an *interpretation* of intellectual processes (that is: of semantic inferences). The way we conceive of 'planning', it can easily be said to be the kernel of 'willing': a conscious choice among alternatives, namely. The 'intensity' of a volition can be interpreted as being the distinctness or reliability of a certain plan's attractiveness and adequacy; a 'weak will' can then be concluded to depend upon the regarded person's inability to discover reliable goals, which can either depend on too little or on too far-reaching reasoning. (A 'strong will' would thus be something similarly ambivalent as 'courage'; 'courage' can depend upon a disregard of certain involved risks, a 'strong will' upon a disregard of certain involved disadvantages, ethical problems, or alternatives. It is typically difficult to tell where courage ends and recklessness and stupidity start, or where a strong will ends and fanatism or obsession (and recklessness once more) start.)

The proverb "The mind is willing, but the flesh is weak" suggests that a distinct intended plan must be innerved also. (Under hypnosis experiments, for instance, we can be told that we now are unable to lift our arm, but we realize clearly that we *could* if we really *wanted* to. If we are curious about hypnosis, we do not really want to 'break the spell', of course. Similarly, we may lie in bed and have the strong intention to rise, but hesitate with the effector innervations that will lead to the kick with which the bed cover is thrown aside and we start to freeze.)

Volition has also to do with plan priorities. Everyone who ever has (truly) participated in a 5000-meter running competition, for instance, knows that it takes 'will' to hang on when other runners increase the pace. One more or less consciously suppresses the reconsiderations of priorities, because one knows from experience that one would regret such a re-valuation as soon as one is back in a 'normal' physiological state.

But 'will' requires also a conviction that what one tries to do is feasible. We are inclined to give up when we realize that our means are insufficient for reaching the intended goal, and if we have no hope that there will be a change in our relative competence. A small sign of fatigue by the competitor can be of essential significance for one's judgment of relative strength, and, in contrast, if a runner seems less exhausted that he really is, it can nevertheless discourage his competitors from 'hanging on'. One's reputation is thus evidently also of great importance in these situations.

8.1. Volition and Ambition

The difference between the meanings of "planning" and "willing", or between "ambition" and "volition", are conventional (cultural). The word "planning" can be used in situations where we make plans for others, while "willing" is not usually used in this sense. The word "ambition" is usually given the meaning of "long-term volition". The implementation of a long-term plan can, of course, become subject to social influences to a far greater extent than the execution of a momentaneous plan ever can be; that is why 'ambition' usually is studied by social psychologists, while 'volition' is said to belong to the realm of individual psychology.

8.2. Free Will

It seems obvious that 'free will' cannot mean a non-causal determination of behavior; random behavior is rather the *opposite* to willed behavior. On the other hand, if we can clearly discern that an individual had no choice, but was *forced* to do what he did, we also say that he did not have a free will (and that he insofar cannot be held responsible for his behavior). Hedenius (1964) points out that we tend to call an act freely willed if a 'normal' person could have chosen to behave differently in the situation in which the studied act was emitted; the actor's psychological constraints upon him must be considered part of the situation description (similary: Frankena 1968, p 59.)

The attribution of free will is socially relevant for attributions of responsibility and motive. Free will is not a psychological function but merely a modelling category used in everyday life. It denotes the fact or hypothesis that a person can (or could) make several alternative plans that involve his own behavior.

(Relevant literature: Steiner 1970.)

9. Self-Awareness and Self Image

Will our simulated robots have self awareness? It seems obvious that they will not, as little as there will be any 'real' emotions when we simulate them, or 'real' robots in the first place. The question must therefore be interpreted thus: can we simulate self awareness in our robots?

We adopt the kernel of our answer from Steinbuch (1963, p 275-276) who points out that the human skin contain some hundred million receptors which define a distinct boundary between the human individual and his environment. Such a boundary is a logical prerequisite for a distinct conceptual representation of Self in the robot's belief system. If we assume that general awareness is a given capacity of our robots, which operationally expresses itself as a capacity to make semantic inferences in the belief system, and if we assume further that emotions and pain sensations are conceptually represented in the belief system, it is obvious that these can be linked into implicationsl molecules. Observe that we must assume that the *conscious* Self *learns* that he tries to satisfy certain basic needs while he already does *subconsciously*. That is: the basic motivational processes come into existence long before the Self has an adequate concept of its own embodiment boundaries. As the robot (or the human individual) learns to know the world, he suddenly discovers *himself* in it, and he studies (and learns) how he 'operates'. This is how conscious egoism is born: The Self considers himself as the most important means for achieving the goals he now has become conscious of, and he uses himself consequently as one of the operators that can be applied in the course of problem solvings. The

switching between 'subjective' and 'objective' self awareness (cf. Duval/Wicklund 1972, p 7ff) is quite expectable from this point of view.

Our simulated robot can evidently have this kind of self awareness if he can reason and plan at all. But there is another aspect to this problem, namely self-image formulation and self-image adequacy. Duval/Wicklund (1972) deal with these aspects in their monograph, but they fail to reveal why Self becomes *ashamed* of not living up to (or being able to live up to) personal or social 'ideals'. It is easily understandable that people can *regret* not being ideal, or that they curse their fate and their handicaps, but why are they ashamed?

The explanation that it is through social influence is an obvious one. One assumes simply that shame is a learned reaction, an emotional signalling of social inadequacy, a rudiment perhaps of a fixed processing pattern that should trigger flight. This hypothesis seems attractive, but we do not really see how shame reactions are reinforced and thus taught and learned.

Assume instead that there has been a more general training to avoid social roles which tend to lower the esteem of significant others for Self if Self plays them. Being positively esteemed by significant others is important for whether the social environment manifests itself as benevolent or as hostile. Now, each time Self discovers that he happens to be playing such an undesirable role, he will try to avoid it. The evasion problem is, as usual, solved consciously and subconsciously, and the subconscious processing implies the shame reaction.

But the individual must also learn to accept criticism. Teenagers are known to be much more sensitive to criticism by significant others than adults are, until they have met norm conflicts sufficiently often to realize that one rarely can please everyone.

Why should people try to make 'good impressions' if it were not in order to gain esteem (and thus to render the social environment increasingly benevolent)? Attempting to make a good impression requires that Self has himself an idea about the values held by the other; when strangers meet, they tend therefore be cautious in giving the other clues regarding their identities, and the revelations will be interdependent. (Similarly: Krappmann 1971, p 36; and Berger/Luckmann 1971, p 45.)

Duval/Wicklund 1972 report on a phenomenon which we all recognize, namely that Self avoids to become self conscious. (Psychoanalysts would claim that Self seeks to 'repress' consciousness of his faults.) Considering that thinking does involve emotions, it is not surprising that Self not only tries to avoid embarrassing situations but even tries to avoid thinking of them. (We suppose that psychotherapy tries to train Self not to be as emotionally responsive to the mere thinking of embarrassments any longer.)

But whether Self finds it embarrassing to observe himself must depend upon the situations in which Self has been made aware of himself. If these situations were most often gratifying, Self will namely not re-experience learned embarrassments but rather pride or contentment.

The relationship between gender roles and self images is problematic. It is very well possible that one gender role gives more opportunities for learning self confidence than another. Furthermore, the greater the diversity of roles for one sex, the less distinctly can a person of that sex be judged with regard to the adequacy of his/her role play, and he/she needs therefore never become intensely embarrassed. The gender-role repertoire of one sex may also tend to be as a whole more 'extraverted' than the other, concentrating more on Self's achievments than on certain of Self's personal qualities, 'ugliness', for instance. (But it is typical that women are 'allowed' to make much more far-reaching *retouches* of looks than men may do.)

(Relevant literature: Bem 1972, Steiner 1970.)

PART II: GENERATIVE MICRO-SOCIOLOGY

10. Schedules

The everyday life of a person is embedded in long-term processes, some of which are biological, some psychological, some social and cultural, some economical and political. Some are very personal processes, such as the process of ageing; some are impersonal processes, for example changes in public opinion. Some are particular for only a few persons, others are common to all people.

We shall elaborate a scheme of invariances that are observable in the lives of people in our contemporary Western societies. We call this scheme 'general life trajectory' because it concerns and comprises the long-term constraints that usually are imposed on us, or that we ourselves impose on our lives. We mean choices such as the choice of education, of employment, of marital partner; or biological processes such as puberty, the climacterium, and ageing in general; or events such as the birth of children or the death of a life companion; or social institutions such as obligatory military service. What is typical for these life-trajectory determinants is that they usually cannot or will not occur at just any time in life (as an illness, or an accident, the death of an aunt, or the first encounter with a friend for life could).

The outline of a general life trajectory has already been presented in Volume 1 (section 5.3) and we shall not repeat it here. We shall instead discuss implementation problems.

When we start the simulation, families have already been created in the datastructure (cf. Volume 3). All ages co-occur and there is therefore no problem

whether we shall start the simulation with the matching of couples or with the births of children.

10.1. Life Phases and Development Schemes

We let LIFE be a jig which is started at conception and which keeps track of Self's personal development. This jig needs only be executed now and then, once a month, for instance; it is an auxiliary jig and has therefore maximal priority and attention value zero; its duration is set to 100 years, but it can be abandoned before that, of course. This jig can take care of the updating of Self's age, and it can also induce Self's death according to certain rules and probabilities. Accidents of random severity can be pushed by it, the consequences of accidents can be modelled as no-action jigs which have certain durations and varying attention values, depending upon the severity of the accident.

The LIFE jig 'pushes' two other jigs, one that keeps track of Self's body growth and health (except accidents), and another which associates Self with appropriate 'life phases'.

The jig BODYGROWTH is auxiliary and simply definable as a function that updates Self's length and weight. The neonate weights about 3500 gram and is about 500 mm long. From infancy until puberty, head growth is slow, trunk growth faster, limb growth fastest, but we disregard these subtleties. BODYGROWTH adds simply about 2 mm to Self's length each day during the first 3-5 months, and about 15 grams in weight. Since we are most interested in *differential* developments, we must observe the nutrition Self receives. Malnutrition during early infancy may namely have profound effect upon Self's intellectual, emotional, and physical development (Watson/Lindgren 1973, p 140), and manifest itself in Self's being smaller in size, less active, less

curious, and permanently intellectually retarded in comparison to normal children.

After this period and for 140-164 months if Self is a boy, 116-152 months if Self is a girl, we increase Self's length 1 mm and the weight by 40-60 grams every sixth day. Then we make these additions every fourth day for 5-6 years, and conclude the development process by making these additions once a month only, for 1-3 years. As yet we do not deal with weight changes after this.

Also the tracking of Self's health is not yet worked out in operational detail. This routine will observe Self's nutrition habits, but also his/her way of life in general. It will trigger illnesses with certain probabilities which depend on Self's own health state, and it will record past treatments.

(This side of the total model of Self can easily be elaborated into a detailed model of how Self's physiology reacts to medications; we thing especially of detailed studies of the compatibility or incompatibility of certain medications in patients with particular health records.)

The other jig we mentioned above as being 'pushed' by LIFE is one that merely does the book-keeping on Self's life phases. We saw above already that BODY-GROWTH was differentiated according to Self's age. The jig LIFEPHASES determines from day to day, or in more random intervals, which life phase Self at present belongs to. Life phase, in turn, can be referred to in jig-execution conditions; we can thus, for instance, differentiate between 'childish' and 'grown-up' behavior.

The boundaries between life phases must, of course, be very fuzzy, and we shall describe shortly how we can achieve such fuzziness. Let us look first at the life-scale subdivisions we have in mind, and assume that these subdivisions are biologically and/or culturally justified:

Life phase	approximate boundaries
prenatal period	-9 - 0 months
birth	0 months
neonatal period	0 - 15 days
infancy 1	15 ds-1 year
infancy 2	1 - 2 years
childhood 1	2 - 4 years
childhood 2	4 - 8 years
preadolescence	8 - 13 years
adolescence 1	13 - 17 years
adolescence 2	17 - 21 years
postadolescence	21 - 26 years
adulthood 1	26 - 35 years
adulthood 2	36 - 50 years
adulthood 3	50 - 70 years
old age 1	70 - 80 years
old age 2	80 -100 years

Each period is now associated with a Gaussian sampling curve such that the probability of a time point's belonging to a certain period p depends on its distance from the middle of p. Let t be a time point between the middles of the periods p and q, and let d be the difference between the middles of p and q. Assume q to be later on the time scale. The probability of T's belonging to p is then (q-t)/d, the probability of its belonging to q is 1-p's probability.

Each life phase can be regarded as imposing a new set of 'tasks' upon Self. Since these tasks can be 'pushed' but are not to be executed before a certain life phase has been reached, we can simulate the to and fro of behavior styles in the transition periods between life phases.

It is not necessary to discuss all of the development 'tasks' of the first ten or eleven life phases because textbooks on the psychology of the child (for instance, Watson/Lindgren 1973) and of adolescence (Andersson/Wallin 1971, Hurlock 1967, or Dreyfus 1972, for example) will describe them in more than sufficient detail. (Cf. also volume 1, section 5.3., for an outline.) We can similarly find monographs on the social psychology of ageing (for instance, Williams et. al. 1963) which describe the emerging 'tasks' for the last two or three life phases. But we know of no developmental studies of adulthood. Adulthood is obviously regarded as a 'normal' state of existence which requires no specific studies. It is true that the adult appears in all studies of careers and organizations, or in studies of housewives, but we must eventually try to compose a more allround picture of him.

But we must perhaps illustrate with some parts of the total trajectory scheme how one can start to delineate constraints and to delimit and order sets of alternatives that apply for Self. Our life-trajectory jigs can be looked upon as hypothese that will presumably be valid also for Self, but which need not be. Self can always be made to break out of the web of expectances we have spun around him little by little. Life phases, development rates, educational patterns, daily work routines, habits, obligations and commitments, goals, the rythms of

physiological needs - they all ensnare Self more or less rigidly.

The following example is written in atman-REC (cf. Volume 3). It shows that life is ultimately a system of trivialities in this activistic perspective.

10.2. Short Example of a Development Scheme

We shall choose education as the topic of our example, which we let remain superficial on purpose in order to preserve its transparency. Anyone who has a wish to be more precise, will easily detect where his improvements could be added. Observe also that not all of the mentioned jigs are or have to be defined when we start to execute 'educationjig'; they can be added later on and are, for the time being, automatically substituted by dummy jigs which are handled by the monitor according to the prescribed jig properties.

```
'educationjig: instance 'timestep years 1,
          'duration perpetual, 'priority 11,
          'attention 0, do if age <4 then nogo;
          if age=4 then timestep(jig):= hours 1;
          suspend-until hour 8;

     local a;
     a:= pushjig (kindergartenjig, letting
          autopush:= list(now, days 1, nil),
          duration:= hours 5);
          suspend jig, years 2;
          abandon a;

     a:= pushjig (grammarschooljig, letting
          autopush:= list(now, days 1, nil),
          duration:= hours 6);
          suspend jig, years 4;
          abandon a;
     a:= pushjig  (highschooljig, letting
          autopush:= list(now, days 1, nil)
          duration:= hours 8);
          suspend jig, years 5;
          abandon a;
```

```
      a:= xorpushjig (collegejig,
           letting autopush:= list(now, days 1,nil),
           duration:= hours 8),
           (apprenticeshipjig,
           letting autopush:= list(now, days 1,nil),
           duration:= hours 8);
           suspend jig, years 4;
           abandon a;
      a:= xorpushjig (universityjig),
                     (employmentjig);

      abandon jig |

      'collegejig: instance 'priority 11, 'attention 0,
      'timestep hours 24,

      do
      if family-income < average-income then
      abandon jig else
      if notshf(education-aspiration,
      college educ-frequ) then abandon jig
      else execute jig;

      local a; if day=6,7 then postpone jig;
      a:= nearest 'college
           travels 'to a, letting priority:= 8;
           attends-lessons 'at a, letting priority:=9,
           attention:= random (30,90)
           returns 'from a, letting priority := 5 |

'universityjig: instance 'priority 7, 'attention 2,
      'timestep hours 24,
      do
      if day= 6,7 then postpone jig;
      local a; a:= nearest 'university;
      travels 'to a, letting priority:= 7
      attends-lessons 'at a, letting priority:= 5
```

```
          attention:= random(50,100)
          returns 'from a, letting priority:= 5 |

'employmentjig: instance 'attention 10,'priority 5,
     do
     local a,b,c;
     a:= available jobs;
     b:= forall c in a build list
          (interest-degree 'for x0(c),
           traveltime 'to geoposition x1(c),
           *the payment x2(c),
           *the firm name x3(c),
           competence 'at x0(c))
     c:= ordering '(1,3,-2,5,nil),b;
     pushjig (apply-for-jobs c; jobjig);
     abandon jig |

'apply-for-jobs: instance 'priority 7, 'attention 20,
     'duration perpetual,
     do
     local a,b; if empty then abandon jig;
     suspend-until hour 8;
     a:= geoposition 'of x3(x);
     b:= traveltime 'to a;
     travels 'to a, letting duration:= d;
     the-very-application 'at x3(x),
          letting attention := 100;
     returns 'to be letting 'duration:= d;
     apply-for-jobs rest |
```

And so on. The very application is a judgment by the firm regarding Self's usefulness for, and fit into, the firm. The firm waits with its decision for five days and compares all persons who have applied during this period.

The jobjig supplies Self with a workdaily curr-

iculum: he travels to the firm in the morning, performs his duties there, returns home, and enjoys his evening leisure finally.

Leisure is much more difficult to model, because it is more 'amorphous', less rigidly persecuted, often subject to situational factors, moods, and interests. Tv-watching, hobbies, housework, conversation, entertainment, games and play, recreation, meeting people, taking a walk, window-shopping, et cetera, compete for Self's attention, and he will give them basic priorities according to his interests. (Cf. Grušin et. al. 1970; Küng 1971.)

Watching television is a good example of a time-triggered event being participated in selectively, depending upon interests, mood, availability of time, and access to certain equipment (here a tv-set).

It is, of course, in principle possible to describe each television program explicitly in terms of interest-relevances; we could then compare program profiles with interests and decide whether profile similarity passes a threshold (which may change with mood and stress).

It may seem doubtful whether such explicit detail is necessary in a non-specialized model. We could instead look at Self's interest profile and estimate how often (on the average, given a certain program policy) Self will be interested to watch a program. We realize on second thought, however, that we must still time-trigger these interesting programs, because we may want to look at the differential effect of tv-watching upon family cohesion when interests are similar or different in a married couple. If a couple watches together and

abstains together, they do not disturb each other, may have something to discuss, accumulate a similar belief base, and have time to converse when they do not watch. On the other hand, if differences in interest lead to asynchronous tv-watching, the differences will be accentuated and positive interaction possibilities be reduced.

10.3. Daily Curricula

Most people have quite regular daily curricula, at least on workdays. For instance:

```
'curriculum1:rise;workhours;evening;gotobed;
          curriculum1 |
'rise:suspend jig, hours 1; breakfast |
'breakfast: suspend jig, hours 1 |
'workhours: for i= 1 to 5 do
          action 9; action 10 |
'action:prstr" performs action "; prin1 x;
          suspend jig, hours 1 |
'evening: action 17; action 18; action 23;
          action 25 |
'gotobed: suspend jig; hours 1; sleeps |
'sleeps: suspend jig; hours 7 |
task(hour 6, person 1, 'curriculum1) |
```

Such curricula suggest (with less than maximal priority so that they can be overridden) how a certain type of person can be expected to spend his or her day. All such curricula are time-steered, but event-triggered behavior can disrupt them at least temporarily.

11. Maps, Places, and Travels

Each physical context is associated with a 'map'. This map is a Cartesian product of positions, some of which are 'occupied' by members of the context. A bus, for example, comprises a number of seats; these seats can be occupied by persons who have entered the bus.

The context itself can 'occupy' a position in the map of another context. To be more precise:iff a context is a member of another context, it 'occupies' a position in the map of that other context. For instance, the above-mentioned bus 'occupies' a position in the map of the village or suburb or town or city where it runs.

Movements within one map do not affect movements in another. The relative positions of the persons seated in a bus are not changed by the bus movement in its supracontext's map, for instance.

11.1. An Example: Travelling by Bus

The following example is again written in atman-REC. It defines two bus routes and simulates two individuals who leave their homes in the morning and take an appropriate bus to the firms where they work.

```
'route1:(if time-of-day<hours 6 then wait hours 6);
        busstop place 1; wait minutes 10;
        busstop place 2; wait minutes 20;
        busstop place 3; wait minutes 5;
        busstop place 4; wait minutes 15;
        busstop place 5; wait minutes 10; route1|

'route2:(if time-of-day<hours 6 then wait hours 6);
        busstop place 1; wait minutes 10;
        busstop place 2; wait minutes 10;
        busstop place 6; wait minutes 5;
        busstop place 7; wait minutes 5; route2|

'busstop:enter x; tell(self.'in,x);
        tell(x.'in, self.'line); leave |
'tell: if x=nil then nil else signal (x0, rest);
        tell(x1, rest) |
'signal:sender:=self; local self; self:=x; execute|
'takebus:wait for x do leave; enter sender;
        wait for y do leave; enter y |
'walk:leave; wait minutes 10; enter x |
task (hours 6, bus 1, 'route1)|
task (hours 6 + minutes 15, bus 2,'route2)|
task (hours 6 + minutes 30, bus 3,'route1)|
task (hours 8, person 1,'(walk place 1; takebus
        (line1, place 4); walk firm 1)) |
task (hours 8, person 2,'(walk place 1; takebus
        (line2; place 6); walk firm 2)) |
'enter:prstr "***"; prin1(self.'name); push "
        " enters "; prin1(x.'name); com "***";
        self.'position:=x;addto(x.'in,self) |
```

```
'leave:prsts "***"; prin1(self.'name);
       prstr " leaves "; prin1(self.'position.'name);
       com "***"; remove (self.'position.'in, self);
       self.'position:= nil |
```

This example models buses very clumsily because it has to specify explicitly how long the buses will take from one bus stop to another. This can be remedied once we have defined how we intend to compute distances between places on different maps. Each map must, of course, have a specific scale, but it can also be given a 'permeability' value which indicates whether it is easy to travel across the map or if we are to assume bad roads and/or many obstacles.

The above example's buses are furthermore in 'suspended animation' while they travel until they arrive at their next stop. This is inadequate in a detailed study because the bus position is not correct in between the stops. Traffic jams et cetera cannot be modelled unless we define how continuous movements (also between maps) are to be simulated.

II 2. Geographical and Economical Information

It is necessary for detailed studies of traffic flows in Uppsala, for example, or of land use and differential yields in an Ethiopian region, to be able to store topographical and geological information along with the context maps (cf. Odell/Preston 1971). Also ownership information needs often be stored in order to be able to foresee the consequences of land-use changes. For instance, economic and technological support to underdeveloped countries can imply the undesirable consequence that landowners replace an army of petty-tenants on their land by modern machines: land use will indeed be more efficient, but at what price? (This is at least what seems to have happened in Ethiopia with Swedish economic aid, according to a recently televised report on the current famine there (March '74)).

We do intent to simulate regional development eventually, from historical times and into the future. Topographical, geological, meteorological and political information will then be necessary for predicting where trade settlements will spring up and florish, for instance, or how extensive cities can become. We think in particular of a model of the tribal migrations in Europe and Northern Africa towards the end of the Roman Empire.

Our map organization is of particular advantage for representing information of geographical and other attributes, because this can be done in a hierarchical manner. Assume A to be a context which comprises B and C as subcontexts, while C again has subcontexts, et cetera. If A now has the value p on a certain variable P (which has been marked to be implicit in the manner we presently describe), we

assume that also all subcontexts of A have the value p on P, *unless* we have explicitly stored other information on P for some subcontext. The subcontexts of that 'deviant' context are then again assumed to have the 'deviant' context's P value, *unless* we have specified some explicit P value for them. Et cetera, down to basic contexts. It is obvious that we thus can derive the P value for each point in the mapped universe without necessarily duplicating very much information in storage.

12. Social Contexts

We have recently come across a book on the 'form and substance of urban social networks' (Laumann 1973) which in a superb manner (although in quite different terms) explicates some of the phenomena and constraints that we have intended to model with the help of the social-context category. We wish also to point to the sociology of Segerstedt (1966) which similarly stresses the importance of social contexts (and which we suspect in retrospect to have influ enced our own thinking towards an explicit representation of social contexts). But also the sociological use of the Marxian class concept or of Weber's 'status group' ("Stand") resemble the use we intend to make of the social-context category.

Social contexts do, of course, not replace more specific categories such as 'class', 'Stand', 'ethnic group', 'peer group', or 'coalition' (to mention only a few). If they are to be used for modelling the specific categories, we must impose specific constraints. That is equivalent to defining in operational detail what we mean if we say "coalition", for instance.

We have described our basic and general view upon the relationship between social contexts and social interaction in some detail in chapter 4 of Volume 1, and we shall not repeat what was said there. We shall instead focus upon a few particularizations of the general view.

First, however, shall we speak about some general characteristics of social contexts.

12.1. Social-Context Membership

All social contexts can be [illegible] with three sets of conditions: entrance conditions, membership conditions, and exit conditions. Entrance conditions specify requirements that one must meet if one wishes to become members of the context (being above 21 years of age, for instance). Membership conditions specify what is required of members if they wish to remain members (having a monthly fee, for example); and exit conditions specify under which conditions a member is allowed to leave the social context (having a passport when leaving the country, for instance). There may also be a special set of conditions for short-term visitors, however.

Very often there are no exit conditions, and the membership conditions may be [illegible] (cf. section 2.7). In fact, we can ground [illegible] Self's membership is a function of [illegible] called 'social acceptance', which we will elaborate below (in section 12.2). The membership conditions can then also be assumed to specify three acceptance-value regions: one of unambiguous membership, one of unambiguous non-membership, and a fuzzy region of ambiguous membership.

We stipulate that the by Self subjectively apperceived 'joining' of a social context C makes only for a weak basic relationship between Self and C. (These apperceptions constitute the 'we feeling' of the members of a social context.) The social context manifests itself as something more than a mere category only to the extent to which its members 'participate' in it, that is: act and interact within it, for it, and according to its norms. Part

of (or implicit in) many social-context norms are therefore demands for participation (for example, going to confession if you wish to be a proper member of the Catholic Church).

12.2. Social Norms and Social Acceptance

The estimation of the social-acceptance value (or measure) is based upon sets of rules which we call 'social norms'. These rules specify criteria which all context members are to follow; the degree to which Self follows them determines to what extent he is approved of in the context. (Observe that these rules do not specify entrance conditions; Self can thus follow the context rules and be highly acceptable, and yet not be a (formal) member. Persons born in one country but living in another exemplify the possibility of such incongruences.)

The social norms have differing priorities which, of course, also affect the social-acceptance measure computation. The degree of Self's compliance with low-priority norms counts usually less than the degree of compatibility of Self and of his behavior with the high-priority norms.

Social norms can be explicit or tacit. We call explicit norms in social contexts 'official goals'. Tacit norms and their priorities must be inductively inferred by Self from observing the behavior of other context members. If the context is new, norms that are valid in similar social contexts will be assumed to hold.

Social norms are guidelines for Self's role play (cf. section 13.1.1).

(Relevant literature: Wright:1971; König 1967; Cohen 1961; Segerstedt 1966; Schwendter 1971.)

12.3. Social Maturation

Also our view on social maturation has already been presented in Volume 1, in section 7.3 namely. Self grows (in Western cultures at least) 'out of' his family of origin and into a multitude of other social contexts (which implies that he obtains rights and duties). As he does so, his autonomy grows, too. Blood (1972, p 284) points out that the child's 'power' within the family increases if he can get moral support from external groups (the school or peers, for instance). William Goode (according to Blood 1972, p 284) has thus more specifically noted that self-selection of a marriage partner (which is an expression of autonomy against the kin) rarely occurs in cultures where adolescents do not form strongly coherent peer groups. But it is not only independence of kin which is fostered in other contexts; independence is quite generally a result of integration into a plurality of independent social contexts (cf. Blood 1972, p 407).

12.4. Acquaintances

Promoters, journalists, politicians and many similar categories of people need to have business 'contacts' in order to be able to be efficient. But the pattern of conditional growth of an 'acquaintance network' that has social contexts as nodes is ubiquitous, not restriced to business only. Looking back at one's past one detects how one involvement led to the next and that few things happened quite by chance. Although we must not overlook that proximity can be a non-social acquaintance-creating factor (cf. Carey/Mapes 1972, p 48ff), it is rare that it leads to more permanent acquaintances, unless it is recurrent.

Acquaintances weight usually also heavier than competences in career competitions. But especially in fields where superiority in competence is difficult to assess, among artists or photographers, for example, we will find that having lots and the 'right' acquaintances ('vitamin A') is crucial for 'getting somewhere'. Observe, however, that obtaining and maintaining the acquaintances also requires certain 'competences'.

13. Imaginary Contexts

The 'social-context' category is an abstraction of a multitude of specific sociological and everyday conceptions, and so is the 'imaginary-context' category. We prefer to say that *roles* are instances of imaginary contexts rather than of social contexts in order to preserve a certain distinction analogous to that between a group and a category. We do namely not expect to be able to compute cohesion measures (cf. Moreno 1953; Hofstätter 1956, p 86f) on imaginary contexts, for instance, while we should be able to do so on social contexts. Also the *reference-group* category can be approximately subsumed in the imaginary-context category if one wishes to detach sets of norms ('subcultures') from specific social contexts. Social *institutions*, however, will be implicitly modelled in social contexts, although "institution", in the ambiguous jargon of sociology, usually refers to a complex of roles, and one therefore should expect institutions to be modelled as imaginary contexts also. But this is not practical: Role differentiation of the members of a social context will instead be modelled such that the various actors belong to different imaginary contexts. This models exactly what "institution" refers to, namely the differentiation and complementarity of roles within a social context.

13.1. Roles

Roles resemble social contexts insofar as the members of these contexts are subject to constraints and obligations. There are thus entrance conditions, membership conditions, and exit conditions, but membership can also be ascribed by others. It is Self's own felt involvement, however, which determines to which degree the imaginary-context membership influences his behavior.

We wish to let role behavior resemble language: there is an underlying semantic structure which can have many different surface manifestations. That is: certain sets of norms are differently interpreted in different social contexts. (This, by the way, is the inverse of the abstraction process through which the role concept was born, when one namely tried to detect behavior invariances in various social contexts.) The roles which we model as imaginary contexts are abstractions, not concrete behavior-sequence prescriptions; the latter are much more easily modelled as jigs (as curricula, for example).

If Self violates the norms of one of his roles, the role's own reaction rules will operate. Self's value on certain variables may, for instance, be decreased, which can affect Self's acceptability in certain social contexts. The visibility of 'role performance' depends on the generality of the variables which are affected by role-inherent sanctions.

(Relevant literature: Dahrendorf 1961; Siegrist 1970; Andersson/Wallin 1971, p 70ff; Biddle/Thomas 1966; Gross et. al. 1958; Segerstedt 1966.)

13.1.1. Role Genesis and Modification

Turner (1962, according to Buckley 1967, p 146ff) has objected to the conception of roles as rigid prescriptions connected with certain poritions in social contexts. He sees roles instead as "tentative processes of reciprocal responding of self and other" (Buckely 1967, p 146).

We agree with Turner and Buckely, but wish to retain loose frames (the imaginary context) which can supply default rules to our simulated agents. Such rules can be specified by us in jig form if they are prescriptive. Restrictions are more difficult to represent; a not very elegant but manageable way is to formulate them as jigs which stop certain other jigs.

The *dynamic* genesis and modification of roles must be modelled on the psychological level. There will always be attempts to gain advantages in interaction and to perpetuate them, but Self must also be assumed to realize that he cannot go too far. He must be able to estimate 'role boundaries' from reactions upon small transgressions during interaction. Self perceives roles as being ways of behavior which are generally accepted and for which he cannot justly be sanctioned negatively, but which are not always easiest or most advantageous for him. If they are not, he will try to modify them (if he dares) by making more or less deviant (but for him more adequate) behavior more frequent and 'normal'.

13.2. Mass Media

The behavior of mass media cannot be immediately influenced. There is little feedback from its average 'consumer' to the medium, and the little there is is delayed. Mass media are modelled as imaginary contexts which overlap with social contexts that influence them. To 'consume' a mass-medium message is to be a current member of the message context. The very message is either a text-interpretation task given to Self, or a conditioning process. (Cf. Goranson 1970, which studies the effects of exposure to depicted violence.)

The types, names, and number of mass media in our model must, of course, be variable and some may even be simulation results. Types could be television, newspapers, books, magazines, radio, and records.

Mass media tend to have physical target contexts (city 1 or region 9 or county 5, for instance), and they can often be said to be 'stationed' there. Within these phsyical contexts they have usually also social target contexts. But the 'reception areas' of mass media must be assumed to be fuzzy with more or less narrow membership functions based upon the distances from the target centers. A single television broadcasting station, for instance, has ordinarily a narrower reception area than a newspaper.

14. Coalition Theory.

A crucial and basic fact of life is that since the world's resources are limited, there is a continuous fight for the control of them. If anyone is to have a servant, there must be someone who does not have one, for example. Flatterings, pledges, persuasions, deceptions, implorations, accusations, commands, threats and violence belong to the repertoire of activities that people emit selectively towards others during the continuous struggle about means and occasions for achieving one's goals. Despite identical (basic and abstract) intrinsic goals, people need not have identical concrete goals because their embodiments and past trainings can have biased them differentially. And even if they have similar concrete goals, people can disagree regarding the means to be employed in order to achieve the goals. And should they all agree on goals and means, but not all can attain the goals (at least not to satisfactory degrees, where degrees are possible), there will still be the disputes about the distribution of 'rewards'. This all goes to show that people must realize that they need to join 'coalitions' of persons with more or less specifically similar goals, and that they need adapt to, and cooperate within, these coalitions. But there will also be disagreements and struggles *within* coalitions, about goals and means and distributions, and there will therefore be subcoalitions ('fractions') except in very small and/or already unanimous coalitions. There will thus be hierarchies of coalitions which all compete for influence, but which still cooperate against such other coalitions which have (or are believed to have) incompatible goals.

Each Self realizes that struggling for a share of the total but limited amount of need- reducing 'resources' is a necessity for survival. He is taught early by his family that one needs the support of coalitions in this struggle; life is no longer only a fight against Nature, nor always a battle of all against all, but very often also a maneuvering of coalitions against coalitions. Self learns that his 'social security' depends upon his being accepted and esteemed by relatively powerful coalitions. His family context, which is the first coalition of which he has been a member, will not suffice to lend him sufficient support and security after he has come to age, and he must therefore obtain membership in other coalitions. Which coalitions will he seek to join?

Self's life procedes in certain physical, social, and imaginary contexts, within which certain coalitions will fight for 'territory'. It is not always necessary to join one of the coalitions (a local religious sect, for example), but sometimes it is.(Self cannot normally live without any citizenship, for instance). Especially in his career it is often important for Self to belong to the 'right' coalition, and to have its promotion.

14.1. Solidarity

A *coalition* is a social context where all members are expected to manifest 'solidarity' with each other (and thus with the coalition). We equate 'solidarity' with 'commitment' or 'dedication'. A person 'Self' is committed to (and thus solidary with) a coalition to the extend to which he

(a) aspires the coalition's official goal and is satisfied with the tacit norms,
(b) has 'invested' time, money, opportunities, or prestige in the coalition,
(c) believes in the success of this coalition,
(d) believes in the stability of this coalition,
(e) believes to be obliged to this coalition by norms and expectations,
(f) believes in being able to obtain at least a fair share of any gain made by this coalition,
(g) is unaware of other favorable coalition opportunities.

We can obtain a solidarity measure by first projecting the values (a) to (f) onto the interval (0,1), then computing the (possibly weighed) average of the projection results for each of the alternatives which Self apperceives, and thereupon computing the measure H (as defined in the appendix to this volume) on the set of these averages; the value 1-H is then a measure of solidarity as specified above.

We can use the solidarity measure for determining whether a given social context is a coalition; let this be a fuzzy classification. But note that the simulation needs not make use of this classification because it can employ Self's solidarity

measure directly in computing Self's decisions.

We can also use the solidarity measure in the course of Self's planning by letting him estimate his hypothetical solidarity if he were a nember. The value (b) must naturally be ignored in such uses.

The value (f) reflects Self's 'social security' within the studied coalition. We shall now take a closer look at this measure.

The condition (g) is tacitly fulfilled through our use of the measure H.

(Relevant literature: Kiesler 1971, more as a contrast than as a support. Also Buckley 1967, p 188-190, and Zander 1971.)

14.2. Attractiveness

Let us say that Self apperceives Other as 'congenial' to the extent to which he believes that he could be solidary with a dyad comprising Self and Other. Such an estimation could be made with the solidarity measure in section 14.1, where past investments have to be ignored, of course.

Let 'attractiveness' be a wider concept than 'congeniality'. Let attractiveness be influenced not only by conscious long-term considerations but also by subconscious and short-term evaluations. Attractiveness can thus be spontaneous while congeniality cannot be assessed at once. After a while, however, can the consciously estimated value outweigh the subconscious evaluation.

14.2.1. Social Apperception

We shall assume that Self apperceives other persons similarly to how he apperceives whole situations, namely first of all subconsciously, and sometimes also consciously. That is, to be more explicit: Self is assumed to comprise a person apperceptor and to learn to classify persons as a consequence of indoctrination, hearsay, and own experience. Some people will thus seem more 'attractive' than other ones, without Self's being able to put the finger to why this is so. The apperceptor input can be a 'looks'-vector and a vector of various other more or less superficial attributes of people.

If we made a factor analysis of the attributes used by Self for such classifications, we would presumably be able to identify factors equivalent to those we mentioned above in connection with the

solidarity measure (except (g)). We assume that three independent dimensions would obtain: *goal compatibility*, Self-complementary *instrumental value*, and '*fairness*'. The last dimension refers simply to Other's probable fairness during intra-coalition distribution processes when a coalition has had success and a 'prize' is to be shared; Self must always be suspicious of being tricked or bullied out of his fair share.

There are two proverbs, to be found in many cultures, which concern person apperception and coalitioning, namely "Contrasts attract one another" and "Birds of a feather flock together". These proverbs are apparently contradictory, but both are popular because they still seem valid. We suggest that if we distinguish between congeniality due to goal compatibility and congeniality due to instrumental value, we can resolve the contradiction: persons who share goals and values have reasons to communicate and to cooperate; persons who are complementary to each other with regard to various instrumental resources (for instance such vaguely instrumental resources as beauty and prestige, or such down-to-earth resources as money and physical strength) must find each other useful, valuable, interesting, and thus 'attractive'. (Similarly: Rosenberg 1956.)

Other's goal compatibility, instrumental value, and fairness can not only be estimated subconsciously, but also from conscious analyses of knowledge and beliefs about Other. The conscious analysis can come to contradict the subconscious evaluation; Self may, for instance, have reason to suspect that an apparently very friendly person is an impostor or a scheming flatterer. Self will try to discover

motives underlying Other's behavior (cf. Jones/Davis 1965), and he will seek to detect invariances in what he knows about Other's past behavior. If Other has been unfair in the past, he will probably be unfair in the future also, for example. The subconscious classification ('first impression') of Other can be used as a hypothesis in this analysis, and intermediate analysis results can evoke stereotype conceptions that are then again used as hypotheses, et cetera.

Coalitions are apperceived similarly to how persons are apperceived. Self starts with evaluating representatives of the coalition and he generalizes his impressions to the whole coalition. He may then analyze the constraints upon this coalition, the relative influence of the representative within the coalition, for instance, in order to find out whether his generalization seems appropriate. Self's image of the coalition will be especially influenced by Self's interpretation of the coalition's past behavior.

Especially here-and-now affiliations, which are not expected to last, are usually not made on the basis of conscious analyses. We presume instead that subconscious situation and person evaluations (and associated emotions) usually suffice for Self for judging whether he can be comfortable, safe, and relaxed in certain short-term company. It is typical that Self is not self-conscious among friends and when he is having a really good time, which indicates that he is not consciously 'scheming' simultaneously. Such affiliations are also usually not called 'coalition', but simply 'group' ('expressive group' more specifically, if coalitions are called 'task group').

14.2.2. Making Impression

Self tries to make 'good impressions' , of course, especially upon those with whom he wishes to co-operate, but also quite generally in order to avoid resistances. What is a 'good impression' will obviously differ from occasion to occasion; it is sometimes to seem very competent and at other times to seem very modest; it is sometimes to seem very active and at other times to seem passive. (Similarly: Blau 1964, p 34-50, p 69ff.)

'Social norms' are explicitly stated or inductively inferred rules which can guide the behavior of social agents by indicating which types of behavior will make 'good impressions' and which 'bad' ones, and also often on whom. Which norms are valid for Self, and with which priorities, depends very much upon his whole set of interests, affiliations, and upon the present context, and cannot be estimated in isolation or in general. One speaks in this connection often of positive or negative 'reference groups' and means something similar to coalitions which one supports or opposes; these 'reference groups' are conceived of as 'norm sources' of varying impacts (cf. Segerstedt 1966, p 10-11, p 77ff; Merton 1968, XI; Berelson/Steiner 1964, p 434-435; "reference group" has unfortunately also another meaning in connection with discussions of 'relative deprivation, cf. Merton 1968, p 40, and X).

14.3. Social Power

Let "social integration of Self in C" denote the degree to which a person 'Self' is integrated into the communication network within the social context C. We can estimate this degree by counting the proportion of all messages sent and received by Self relative to the number of all messages sent and received by anyone in C.

Let "social importance of Self in C" denote the proportion of messages which mention Self, compared with all messages sent in C.

Social integration, social importance, and social power tend to be strongly positively correlated, and should therefore be usable as symptoms of one another. But social integration and social importance are inappropriate as the bases of a definition of social power, because it is difficult to explain how they have come about if one cannot refer to social power.

Let us therefore assume quite different, originally non-social bases, such as competence, physical strength, material wealth, or beauty; but also merely subconsciously derived 'attractiveness' estimates can be used here for characterizing a certain person 'Self'. Let "A(Self,U,C)" denote the result of the person apperception of Self performed by the member U of the coalition C. Let consequently "E (Self, C)" denote the sum of the person-apperception results A(Self,U,C) for all U in C, except Self. This sum be called 'C's esteem for Self'.

Compute next C's esteem for each ('M') of the members in C, and then normalize each esteem-value E(M,C) by dividing it by $\Sigma E(M,C)$, the sum of all

esteem-values computed in C.

These normalized values can be said to reflect each M's *relative power in the coalition C*. But we shall use these values only as first approximations. We wish namely to take into consideration that Self's social power is highly dependent upon the degree to which Self can expect to be *backed up by*, or at least compatible with, other persons or coalitions. That is: we wish to let Self's relative power in C depend upon Self's relative powers in all the other coalitions in which Self is currently found to be a member.

We compute therefore the first approximations of the the relative powers of *all* members in all coalitions in which any member M of C is a member, accumulate a sum P of relative-power values for each M, and compute then a sum Q of these sums for each coalition. After having normalized all sums Q by dividing them by ΣQ, we can use them as weights S for recomputing the sums P. The idea is to let the 'status' of a coalition weigh the contribution M receives from being a member in it. Normalizing P anew gives us approximations of the relative powers of the members of C.

We can now compute second approximations of relative powers for all members of any referred-to coalition and thus eventually arrive at third approximations, if one desires. Higher-order approximations can also be obtained in this recursive manner, but can hardly improve the measure's adequacy. We find namely that each approximation creates greater consensus about the judgment of relative power, and we believe that we must not let such consensus become overwhelming.

Let us finally define an absolute social-power measure as the sum of all the relative powers of Self in the coalitions in which he currently is found to be a member.

It is often much easier to overlook such computational procedures in formalized form:

Step 0: Preliminaries: Let n be an approximation counter. Start with n=1.

step 1: Assume A(M,U,C) for all M and U in C. We may also assume that not all U know all M although they belong to the same coalition; let their default appercepts be 0.5, which implies that being unknown can be better than being disliked; let the default appercept be 0.3 if U is to be suspicious against strangers.

step 2: $E(M,C) := \Sigma A(M,U,C)$.

step 3: $Ebase(C) := \Sigma E(M,C)$.

step 4: $R_n(M,C) := E(M,C)/Ebase(C)$.
This is the nth approximation of the relative power of M in C, to be computed for all M in C.

step 5: Let $P_n(M) := R_n(M,K)$ for all K in which M is currently found to be a member. This is the nth approximation of M's total and absolute social power within the universe Z of coalitions which we take into consideration. $P_n(M)$ is to be computed for each M in Z.

step 6: Let $Q_n(C) := \Sigma P_n(M|M \text{ in } C)$. This is the total and absolute social power of the coalition C, based upon the P_n-values of its members.

step 7: Let $S_n(C) := Q_n(C) / \Sigma\, Q_n(K \mid K \text{ in } Z)$. This is the relative social power of C in Z; compute $S_n(C)$ for all C in Z.

step 8: n:= n+1; m:= n-1. We can now improve our last approximations.

step 9: Let $P_n(M) := \Sigma\, R_m(M,C) \cdot S_m(C)$, for all C in which M currently is found to be a member. We let the new absolute social power of M depend on M's status in C and C's status in Z. M's membership in C contributes maximally to his total social power if he is powerful in a powerful coalition.

step 10: Compute step 6.

step 11: Let $R_n(M,C) := P_n(M)/Q_n(C)$. This is the new relative social power of M in C.

step 12: If n<nmax then go to step 7, else stop. "nmax" denotes a given maximal order of approximations, 2 or 3, for instance.

By defining the measure recursively, it can be kept mathematically trivial in each detail, as one sees.

14.3.1. Compatibility

We have assumed for our social-power measure that the degree of Self's social power in a coalition also reflects Self's compatibility with the goals of that coalition. The social-power measure is to be used during Self's and Other's planning and making choices, it does not trigger action directly. But how it can influence the choice of action shall be elaborated somewhat.

Self experiences the regulating efforts of other persons sometimes as assistence, sometimes as

resistance. Support and resistance can be directed towards Self on purpose or by coincidence. Purposive assistance of, or resistance against, Self is emitted by others for the same reasons for which Self supports or resists others.

Self supports such coalitions which seek to implement Self's goals or subgoals of these. This support needs not always be tangible, it is perhaps even seldom so, but it can express itself in positive 'attitudes' if one provokes Self to pass judgment. Active cooperative involvement in the work of a coalition requires an impetus: Self has either for some reason become extraordinary motivated to actively strive for the coalition goal, or he has been introduced by another person into the coalition and admonished to cooperate.

Self is in opposition to such coalitions which seek to prevent the implementation of goals which Self supports, and/or which seek to implement goals which Self wishes to prevent from being attained. Like support, opposition needs not always be tangible; Self will assign priorities to support tasks as well as to resistance tasks, and only the highest-priority regulation tasks will be actively tackled.

If he is consistent, Self must also support opponents to his opponents, support supporters of those he supports, oppose supporters (and supporters of supporters) of opponents, et cetera. It is impossible, of course, for Self to know about each and every other person whether he shall support or oppose him; besides, indirect support and indirect opposition will probably seem to demand lower regulation priorities.

Nor must we believe that compatibility or incompatibility is clear-cut. Coalition P or person P may be compatible with Self with regard to one question, but incompatible with regard to another. Self aspires namely not only a single goal but a whole set of goals, and these goals can again differ in priorities.

The *concrete* estimations of compatibility and triggerings of supporting or inhibiting actions cannot be discussed on the level on which we presently speak about these things. Such concrete discussion needs namely to relate to a concrete belief system. A *measure* of the net of social support and resistance apperceived by Self would also have to be based upon the belief system. This measure is estimated by the measure P of Self's absolute social power (in section 14.3 above).

15. Emergent Properties

We have now specified the essentials of our generative micro-sociology and shall next comment on a few epiphenomena (such as leadership) and surface phenomena (such as attitude-balancing) which often are found to be central in textbooks on social psychology.

15.1. Attitudes

Attitudes are consequences of social-resistence analyses (partially similar: Katz 1960). The word "attitude" is commonly used to denote "certain regularities in feelings, thoughts, and predispositions to act on the part of an individual toward some aspect of his environment" (Secord/Backman 1964, p 95). It is obvious from what we have said earlier that attitudes can be subconsciously and/or consciously arrived at, and they can therefore comprise affective and/or cognitive components (cf. Rosenberg 1956, but observe the difference in denotation). When these components are mutually inconsistent, Self can *only by conscious analysis* try to resolve the inconsistency (which does not exclude cases where the subconscious evaluation changes as a consequence of conscious belief changes, or due to new experiences). Inconsistency resolution is not guaranteed by such analysis, unless Self is led to consciously change at least one of the three evaluation factors (or Rosenberg's two) we mentioned in section 12.5.1. Note that we, in contrast to Rosenberg, consider all these evaluation factors to be *cognitive* dimensions.

15.1.1. Cognitive Dissonance

It is obvious to everyone that if Self believes that p and not-p cannot be true simultaneously, and he is told that q and not-q is true simultaneously, he w will start to investigate which of the statements needs to be modified or differentiated, if he is at all interested in the subject matter referred to by the statements (cf. Volume 2, sections 9.2 and 9.4.3). This interest will by itself probably be low if the

fuzzy validities of the inconsistent statements are close to 0.5 (cf. Festinger's notion of dissonance magnitude in this connection).

15.1.2. Resistance to Attitude Change

We assume that resistance to attitude change depends upon how 'well-founded' the attitude is. If it is based upon fundamental values, reliable data, and simple logical inferences, Self cannot change his attitude without evoking worse dissonances. In modelling Self, we need not compute estimates of probable resistance to attitude change in order to decide whether we may make a certain change; a belief system modelled instead according to the principles discussed in Volume 2, section 9.4 and chapter 10, will 'resist' or 'comply' with perfect adequacy.

Self's total persuasibility depends upon how 'solid' his world view is (in terms of comprehensiveness, coherence, and consistency).

15.1.3. Freedom of Choice and Dissonance

Secord/Backman (1964, p 156-159) report that the magnitude of dissonance increase in Self with regard to an action choice seems to depend upon Self's perceived freedom of choice. We suggest that "dissonance" is here to be replaced by "uncertainty" which has a more general meaning: the more uncertain Self is of how to choose among several alternatives, the greater is his perceived freedom; the reverse is true by induction. This is consistent with what we said in chaper 8 on volition and 'free will'

15.2. Socialization (Social Learning)

Social learning is distinguished from non-social learning to the extent to which it concerns *conventions* such as names, rituals, norms, roles, or modes of thinking. We do not stipulate that such learning needs occur during immediate social interaction. Secord/Backman (1964, p 526) point out justly that "socialization should not be thought of as molding a person to a standard social pattern, however".

Self's social learning is constrained by his past and present social-context memberships, of course. The order of making acquaintance with certain contexts can have profound importance for consequent resistance to attitude and belief changes.

The *simulation* of social learning is easier than that of non-social learning insofar as training instructions explicitly point out the 'correct' responses. There is no other difference. The problem of socialization is thus a problem of process *interpretation*, not of process *generation*.

15.3. Group Dynamics

Groups are social contexts, but not all social contexts are groups. Coalitions are groups, but not all groups are coalitions. That is: 'coalition' ⊂ 'group' ⊂ 'social context'. Social contexts are called 'groups' to the extent to which the interaction between the context members is 'dense'. Interaction density is a relative measure: It is the proportion of attention dedicated to other context members as compared with the attention dedicated to non-members (cf. Hofstätter 1959, p 309-311, p 319).

Groups are called 'coalitions' to the extent to which they are 'cohesive'. Cohesion is a relative measure of solidarity: It is the sum of the factual solidarity measures of all members compared with the maximal possible sum.

15.3.1. Leadership

Much has been speculated about what makes a leader a leader (cf. Hofstätter 1959, p 339ff). Note that such speculations are not made unimportant by our social-power measure; we cannot simply say that the leader of a coalition C is the person with the maximal relative social power R_1 in C. R_1 does not allow us to predict which persons will 'advance' and which 'retard' in the ensuing positional warfare. Personal attributes such as 'pondus', 'extroversion', verbality, voice, or experience will play an important role, in part because they render a person immediately attractive or unattractive, in part also because they imply a disposition for being active or inactive (communicative, for example) and thus increase visibility, orientation, occasions for making 'propaganda' for oneself, and the probability

of being present (and able to have a say) when sub-contexts happen to make decisions.

15.4. Conclusion

It is certainly evident now how a generative micro-sociology 'implies' ordinary social psychology, and that the latter analyzes and describes processes, while the former generates them. It should also be visible that social psychology still does not deal with pseudoproblems, although the confusion of generative and interpretative ambitions may have distorted the importance attributed to certain topics in the field (dissonance theory, for instance). It is still necessary, namely, to analyze and describe the processes generated by simulation. This does not mean, in reverse, that generative micro-sociology is reduntant because it cannot *replace* social psychology. We can now at last *distinguish* neatly between generative and interpretative aspects and 'unmix' the two into much more concisely defined and complementary sets of problems.

Appendix

The Equidistribution Measure H

Given be a set L of n elements m. Each m is associated with a specific weight p(m), a probability, for instance. Normalize these weights such that

$p/m := p(m)/\Sigma p(m)$, for all m in L.

The measure H is of use in all situations where one seeks a measure which fulfills the following conditions:

(a) H(L) is a continuous function of the weights p(m);

(b) H(L) is minimal if one p(m) equals 1;

(c) H(L) is maximal (for a fixed n) if all m have the same weight; that is, if each p(m)=1/n;

(d) H(L) increases monotonously if n increases.

Shannon (1948) has shown that

$$H(L) = -\sum^{n} p(m)\,{}^{2}\log\, p(m)$$

is a measure which can fulfill these four conditions simultaneously. (Cf. also Khinchin 1957, p2ff; Klír/Valach 1967, p 61; Klix 1971, p 59ff.) Its maximum is ${}^{2}\log n$, its minimum zero.

16. References

Abelson/Reich 1969
R.P. Abelson/C.M. Reich: Implicational Modules: a Method for Extracting Meaning from Input Sentences. *In*: IJCAI69, p 641-647. (Re: 6.3.1.)

Andersson/Wallin 1971
Bengt-Erik Andersson/Erik Wallin: *Tonåringarna och Omvärlden*. Stockholm: Almqvist & Wiksell, 1971. (Re: 10.1; 13.1.)

Ashby 1956
W. Ross Ashby: *An Introduction to Cybernetics*. London: Chapman, 1956. (Re: 2.1.)

Bem 1972
Daryl J. Bem: Self-Perception Theory. *In*: Berkowitz 1972. (Re: 9.0.)

Berelson/Steiner 1964
Bernard Berelson/Gary A Steiner: *Human Behavior*. New York: Harcourt, Brace & World, 1964. (Re: 14.2.2.)

Berger/Luckmann 1971
Peter L. Berger/Thomas Luckmann: *The Social Construction of Reality*. Harmondsworth: Penguin Books, 1971. (Re: 9.0.)

Berkowitz 1965
Leonard Berkowitz (Ed.): *Advances in Experimental Social Psychology, Vol 2*. New York: Academic Press, 1965.

Berkowitz 1970
Leonard Berkowitz (Ed.): *Advances in Experimental Social Psychology, Vol 5*. New York: Academic Press, 1970.

Berkowitz 1972
Leonard Berkowitz (Ed.): *Advances in Experimental Social Psychology, Vol 6.* New York: Academic Press, 1972.

Biddle/Thomas 1966
B.J. Biddle/E.J. Thomas: *Role Theory: Concepts and Research.* New York: Wiley, 1966. (Re: 13.1.)

Bindra 1970
Dalbir Bindra: Emotion and Behavior Theory: Current Research in Historical Perspective. *In*: Black 1970. (Re: 7.0.)

Black 1970
Perry Black (Ed.): *Physiological Correlates of Emotion.* London: Academic Press, 1970.

Blau 1964
Peter M. Blau: *Exchange and Power in Social Life.* New York: Wiley, 1964. (Re: 14.2.2.)

Blood 1972
Robert O. Blood: *The Family.*
New York: The Free Press, 1972. (Re: 12.3.)

Bridges 1932
Katherine M.B. Bridges: Emotional Development in Early Infancy. *In*: *Child Development*, 1932, 2, 214-341. (A summary of findings can be found in Watson/Lindgren 1973, p 155-156.) (Re: 7.0.)

Buckley 1967
Walter Buckley: *Sociology and Modern Systems Theory.* Englewood Cliffs, N.J.: Prentice-Hall, 1967. (Re: 13.1.1; 14.1.)

Carey/Mapes 1972
Lynette Carey/Roy Mapes: *The Sociology of Planning.* London: B.T. Batsford, 1972. (Re: 12.4.)

Cheek 1959
D. Cheek: Unconscious Perception of Meaningful Sounds During Surgical Anæstesia as Revealed Under Hypnosis. *In*: *Amer. J. Clin. Hypnosis, U/101*; 1959. (Re: 2.5.1.)

Cherry 1953
E.C. Cherry: Some Experiments on the Recognition of Speech, With One and With Two Ears. *In*: *J. acoust. Soc. Amer.*, 1953, *25*, 975-979. (Re: 2.5.1.)

Cohen 1961
Albert K. Cohen: *Kriminelle Jugend*.(Orig. title: *Delinquent Boys. The Culture of the Gang.*) Reinbeck bei Hamburg: Rowohlt, 1961. (Re: 12.2.)

Colby 1971
K.M. Colby/S. Weber/F.D. Hilf: Artificial Paranoia. *In*: *Artificial Intelligence*, 1971, *2*, 1-25. (Re: 6.0.)

Coles 1969
L.S. Coles: Talking With a Robot in English. *In*: IJCAI69, p 587-596. (Re: 6.3.)

Dahrendorf 1961
Ralf Dahrendorf: *Homo Sociologicus*.Köln/Opladen: Westdeutscher Verlag, 1961. (Re: 13.1.)

Davitz 1969
Joel R. Davitz: *The Language of Emotion*.
New York: Academic Press, 1969. (Re: 7.0.)

Dorsch 1959
Friedrisch Dorsch: *Psychologisches Wörterbuch*.
Bern: Huber, 1959. (Re: 8.0.)

Dreyfus 1972
Edward A. Dreyfus: *Youth: Search for Meaning*.
Columbus, Ohio: Charles E. Merrill, 1972. (Re: 10.1.)

Duval/Wicklund 1972
Shelley Duval/Robert A. Wicklund: *A Theory of Objective Self Awareness*. New York: Academic Press, 1972. (Re: 9.0.)

Eysenck/Meili/Arnold 1972
H.J. Eysenck/ W. Arnold/ R. Meili (Eds.): *Encyclopedia of Psychology, Vols 1-3*. London: Search Press, 1972. (Re: 5.3; 8.0.)

Field 1960
J. Field (Ed.): *Handbook of Physiology*. Washington: American Physiological Society, 1960.

Fikes/Nilsson 1971
Richard E. Fikes/ Nils J. Nilsson: STRIPS: A New Approach to the Application of Theorem Proving to Problem Solving. *In:* IJCAI71, p 608-720. (Re: 6.3; 6.3.2.)

Findler/Chen 1971
Nicholas V. Findler/David Chen: On the Problems of Time, Retrieval of Temporal Relations, Causality, and Co-Existence. *In*: IJCAI71, p 531-545. (Re: 6.2.)

Frank 1964
Helmar Frank: *Kybernetische Analysen Subjektiver Sachverhalte*. Quickborn bei Hamburg: Schnelle, 1964. (Re: 2.5.2.)

Frankena 1968
William K. Frankena: *Ethics*.
Englewood Cliffs, N.J.:Prentice-Hall, 1968. (Re:8.2.)

Freeman/Newell 1971
P. Freeman/ A. Newell: A Model for Functional Reasoning in Design. *In:* IJCAI71, p 621-640. (Re: 6.3.2.)

Garvin 1970
Paul L. Garvin (Ed.): *Cognition: A Multiple View*.
New York: Spartan, 1970.

George 1970
Frank George: *Models of Thinking.*
London: George Allen and Unwin, 1970. (Re: 3.0.)

Gerard 1960
R. Gerard: Neurophysiology. An Integration.
In: Field 1960, p 1919. (Re: 2.5.)

Goranson 1970
Richard E. Goranson: Media Violence and Aggressive Behavior. *In*: Berkowitz 1970. (Re: 13.2.)

Gross et. al. 1958
N. Gross/W.S. Mason/ A.W. McEachern: *Explorations in Role Analysis.* New York: Wiley, 1958. (Re: 13.1.)

Grey/Wedderburn 1960
J.A. Grey/A.A.I. Wedderburn: Grouping Strategies with Simultaneous Stimuli. *In*: *Quart. J. exp. Psychol.* 1960, *12*, 180-184. (Re: 2.5.1.)

Grice 1948
G.R. Grice: The Relation of Secondary Reinforcement to Delayed Reward in Visual Discrimination Learning. *In*: *J. exp. Psychol.* 1948, *38*, 1-16. (Re: 3.1.)

Grušin et. al. 1970
Boris Grušin et. al.: *Die freie Zeit als Problem.*
Berlin Verlag, 1970. (Re: 10.2.)

Hedenius 1964
Ingemar Hedenius: *Fem Studier i Praktisk Filosofi.*
Uppsala: Filos. Fören., 1964. (Re: 8.2.)

Hilgard /Bower 1966
E.R. Hilgard/ G.H. Bower: *Theories of Learning, 3rd ed.* New York: Appleton, 1966. (Re: 3.1.)

Hofstätter 1956
Peter R. Hofstätter: *Sozialpsychologie.*
Berlin: Walter de Gruyter, 1956. (Re: 13.0.)

Hofstätter 1957
Peter R. Hofstätter (Ed.): *Psychologie*. Frankfurt a.M.: Fisher, 1956. (Re: 8.0.)

Hofstätter 1969
Peter R. Hofstätter: *Einführung in die Sozialpsychologie*. Stuttgart: Alfred Kroner, 1959. (Re: 15.3; 15.3.1.)

Holst/Mittelstædt 1950
E. von Holst/H. Mittelstædt: Das Reafferenzprinzip. *In*: *Naturwiss. 37*, 1950. (Re: 2.5.)

Hull 1937
C.L. Hull: Mind, Mechanism, and Adaptive Behavior. *In*: *Psychol. Rev.*, 1937, *44*, 1-32. (Re: 3.1.)

Hurlock 1967
Elisabeth B. Hurlock: *Adolescent Development 3.ed.* New York: McGraw-Hill, 1967. (Re: 10.1.)

Hutchins 1971
W.J. Hutchins: *The Generation of Syntactical Structures from a Semantic Base*. Amsterdam: North-Holland, 1971 (Re: 6.0.)

IJCAI69
D.E. Walker/L.M. Norton (Eds.): *Proceedings of the International Joint Conference on Artificial Intelligence*. Bedford, Mass., 1969.

IJCAI71
Advance papers of the Second International Joint Conference on Artificial Intelligence. London, 1971.

IJCAI73
Advance papers of the Third International Joint Conference on Artificial Intelligence. Stanford, Calif.: SRI Publ. Dept., 1973.

Izard 1972
Caroll E. Izard: *Patterns of Emotions.*
New York: Academic Press, 1972. (Re: 7.0.)

Jones/Davis 1965
E.E. Jones/ K.E. Davis: From Acts to Dispositions: The Attribution Process in Person Perception. *In:* Berkowitz 1965. (Re: 6.3.1; 14.2.1.)

Katz 1960
D. Katz: The Functional Approach to the Study of Attitude Change. *In: Public Opinion Quarterly*, 1960, *24*, 163-204. (Re: 15.1.)

Keiler 1970
Peter Keiler: *Wollen und Wert.*
Berlin: Walter de Gruyter, 1970. (Re: 8.0.)

Khinchin 1957.
A.I. Khinchin: *Mathematical Foundations of Information Theory.* New York: Dover, 1957.
(Re: Appendix.)

Kiesler 1971
Charles A. Kiesler: *The Psychology of Commitment.*
New York: Academic Press, 1971. (Re: 14.1.)

Klein et. al. 1973
Sheldon Klein/J.F. Aeschlimann/D.F. Balsiger, et.al. Automatic Novel Writing: A Status Report. Technical Report nr. 186, Computer Science Dept., Univ. of Wisconsin, July 1973. (Mimeo.) (Re: 6.1.)

Klír/Valach 1967
Jiří Klír/Miroslav Valach: *Cybernetic Modelling.*
London:Iliffe Books, 1967. (Re: Appendix.)

Klix 1971
Friedhart Klix: *Information und Verhalten.*
Berlin: VEB Deutscher Verlag d. Wiss., 1971.
(Re: Appendix.)

König 1967
René König: *Kleider und Leute. Zur Soziologie der Mode*. Frankfurt a.M.: Fischer, 1967. (Re: 12.2.)

Krappmann 1971
Lothar Krappmann: *Soziologische Dimensionen der Identität*. Stuttgart: Klett, 1971. (Re. 9.0.)

Küng 1971
Emil Küng: *Arbeit und Freizeit*.
Tübingen: J.C.B. Mohr, 1971. (Re: 10.2.)

Lakoff 1972
George Lakoff: Linguistics and Natural Logic.
In: Davidson/Harman 1972, p 545-665. (Re: 6.1.)

Lamb 1970
Sydney M. Lamb: Linguistic and Cognitive Networks.
In: Garvin 1970, p 195-226. (Re. 6.1.)

Laumann 1973
Edward O. Laumann: *Bonds of Pluralism: The Form and Substance of Urban Social Networks*. New York: Wiley-Interscience, 1973. (Re: 12.0.)

Leech 1969
Geoffrey N. Leech: *Towards a Semantic Description of English*. London: Longman, 1969. (Re: 6.0; 6.1; 6.2.)

Marx 1970
Melvin H. Marx (Ed.): *Learning: Theories*.
New York: MacMillan, 1970. (Re: 3.1.)

McCarthy/Hayes 1969
John McCarthy/Patrick Hayes: Some Philosophical Problems from the Standpoint of Artificial Intelligence. *In*: *Machine Intelligence, vol. 4*, 1969, 463-504. (Re: 6.3.2.)

Merton 1968
Robert K. Merton: *Social Theory and Social Structure, 1968 enlarged ed.* New York: The Free Press, 1968. (Re: 14.2.2.)

Mesarovič 1965
Mihail D. Mesarovič: Toward a Formal Theory of Problem Solving. *In:* Sass/Wilkinson 1965. (Re: 5.2.)

Miller 1969
Paul R. Miller: *Sense and Symbol.*
London: Staples, 1969. (Re: 2.5.1; 8.0.)

Moreno 1953
J.L. Moreno: *Who Shall Survive?*
Washington, D.C., 1934. (Re: 13.0.)

Murray 1964
Edward J. Murray: *Motivation and Emotion.*
Englewood Cliffs, N.J.: Prentice-Hall, 1964.
(Re: 7.0.)

Nilsson 1965
Nils J. Nilsson: *Learning Machines.*
New York: McGraw-Hill, 1965. (Re: 4.1; 4.4.)

Norman 1969
Donald A. Norman: *Memory and Attention.*
New York: Wiley, 1969. (Re: 2.5.1.)

Odell/Preston 1973
Peter R. Odell/David A. Preston: *Economies and Societies in Latin America: A Geographical Interpretation.* London: Wiley-Interscience, 1973.
(Re: 11.2.)

Perin 1943
C.T. Perin: A Quantitative Investigation of the Delay-of-Reinforcement Gradient. *In: J. exp. Psychol.* 1943,*32*, 37-51, 95-109. (Re: 3.1.)

Perkins 1947
C.C. Perkins, Jr: The Relation of Secondary Rewards to Gradients of Reinforcement. *In*: *J. exp. Psychol.* 1947, *37*, 377-392. (Re: 3.1.)

Razran 1965
G. Razran. Evolutionary Psychology: Levels of Learning - and Perception and Thinking. *In*: Wolman and Nagel 1965. (Re: 3.1.)

Riesen 1940
A.H. Riesen: Delayed Reward in Discrimination Learning by Chimpanzes. *In*: *Comp. Psychol. Monogr.* 1940, *15*, (5, Whole No. 77). (Re: 3.1.)

Rose 1962
Arnold Rose (Ed.): *Human Behavior and Social Process: An Interactionist Approach.* Boston: Houghton-Mifflin, 1962.

Rosenberg 1956
Morris J. Rosenberg: Cognitive Structure and Attitudinal Affect. *In J. abnorm. soc. Psychol.* 1956, *53*, 367-372. (Re: 14,2,1: 15.1.)

Sandewall 1972
Eric Sandewall: PCF-2, A First-Order Calculus for Expressing Conceptual Information. Uppsala (Sweden): Uppsala University Dept. of Computer Sciences, 1972. (Mimeo.) (Re: 6.0; 6.1; 6.2.)

Sass/Wilkinson 1965
M.Sass/ W. Wilkinson: *Computer Augmentation of Human Reasoning.* Washington: 1965.

Schanck et. al. 1973
R.C. Schanck/N. Goldman/ R.C. Rieger III/ C. Riesbeck: MARGIE: Memory, Analysis, Response Generation and Inference on English: *In*: IJCAI73, p 255-261. (Re: 6.0; 6.1.)

Schmidt/Addiamo 1973
C.F. Schmidt/J. D'Addamio: A Model of the Common-Sense Theory of Intentional Personal Causation. *In*: IJCAI73, p 465-471. (Re: 6.3.1.)

Schwendter 1971
Rolf Schwendter: *Theorie der Subkultur.* Köln: Kiepenheuer, 1971. (Re: 12.2.)

Secord/Backman 1964
Paul F. Secord/ Carl W. Backman: *Social Psychology.* New York: McGraw-Hill, 1964. (Re: 15.1; 15.1.3; 15.2.)

Segerstedt 1966
Torgny T. Segerstedt: *The Nature of Social Reality.* Stockholm: Bonniers, 1966. (Re: 12.0; 12.2; 13.1; 14.2.2.)

Seward 1970
John P. Seward: Conditioning Theory. *In*: Marx 1970, p 49-117. (Re: 3.1.)

Siegrist 1970
Johannes Siegrist: *Das Consensus-Modell.* Stuttgart: Enke, 1970. (Re: 13.1.)

Siklóssy 1972
Laurent Siklóssy: Natural Language Learning by Computer. *In*: Simon and Siklóssy 1972, p 288-328. (Re: 6.0; 6.1.)

Simon 1957
Herbert A. Simon: *Models of Man.* New York: Wiley, 1957. (Re: 1.0.)

Simon and Siklóssy 1972
Herbert A. Simon and Laurent Siklóssy (Eds.): *Representation and Meaning.* Englewood Cliffs, N.J.: Prentice-Hall.

Steinbuch 1963
Karl Steinbuch: *Automat und Mensch, 2nd ed.* Berlin: Springer. (Re: 4.4; 9.0.)

Steiner 1970
Ivan D. Steiner: Perceived Freedom. *In*: Berkowitz 1970, p 187-249. (Re: 6.3.1; 8.2; 9.0.)

Turner 1962
Ralph H. Turner: Role Taking Process Versus Conformity. *In*: Rose 1962. (Re: 13.1.1.)

Uhr 1973
Leonard Uhr: *Pattern Recognition, Learning, & thought.* Englewood Cliffs, N.J.: Prentice-Hall, 1973. (Re: 1.0.)

Watson/Lindgren 1973
Robert I. Watson/ Henry Clay Lindgren: *Psychology of the Child, 3rd ed.* New York: Wiley, 1973. (Re: 10.1.)

Wilks 1973
Yorick Wilks: Understanding Without Proofs. *In*: IJCAI73, p 260-277. (Re: 6.0; 6.1.)

Williams et.al. 1963
R.H. Williams/C. Tibbitts/ W. Donahue (Eds.): *Processes of Aging, I and II:* New York, 1963. (Re: 10.1.)

Winograd 1972
Terry Winograd: *Understanding Natural Language.* Edinburg at the University Press, 1972. (Re: 6.3.)

Wolfe 1934
J.B. Wolfe: The Effect of Delayed Reward upon Learning in the White Rat. *In*: *J. comp. Psychol.*, 1934, *17*, 1-21. (Re: 3.1.)

Wolman/Nagel 1965
B.B. Wolman/E. Nagel (Eds.): *Scientific Psychology.*
New York: Basic Books, 1965.

Wright 1971
Derek Wright: *The Psychology of Moral Behavior.*
Hammondsworth: Penguin Books, 1971. (Re: 12.2.)

Zander 1971
Alvin Zander: *Motives and Goals in Groups.*
New York: Academic Press, 1971. (Re: 14.1.)

Sociocybernetics

is a way to model social reality, and man in particular. It is an integrating interdisciplinary approach which explicitly acknowledges the fact that scientific capacity has been essentially augmented through the advent of the computer as an easily accessible tool. Sociocybernetics can be said to be a modern and operational epistemology for the social and behavioral sciences.

Homo Cyberneticus

is an exemplification of how sociocybernetics is to be used for comprehensive causal modelling. *Homo Cyberneticus* seeks to outline a coherent map of everyday life, upon which computer simulations of social-interaction processes can be based. It can also be looked upon as a research program that can systematize and unify research efforts in otherwise quite dissimilar fields, for instance psychology of religion and robotology, linguistics and ergonomics, or literature and futurology. It seeks to thus reassemble man into one whole.

Homo Cyberneticus comprises 5 booklets:

Volume 1: The Simulation of Everyday Life: A Synopsis
Volume 2: Towards an Epistemology of Dynamic Causal Modelling
Volume 3: Heuristic Simulation Programming
Volume 4: Artificial Psychology and Generative Micro-Sociology
Volume 5: A Macro-Sociological Framework

Please send your orders to your booksellers or to
SOCIOGRAFICA
Vikingagatan 50, S—753 24 UPPSALA, SWEDEN